

ІНСТИТУТ МОЛЕКУЛЯРНОЇ БІОЛОГІЇ ТА ГЕНЕТИКИ  
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ

Кваліфікаційна наукова праця  
На правах рукопису

ЛИХЕНКО ОЛЕКСАНДР КОСТЯНТИНОВИЧ

УДК 577.218

## ДИСЕРТАЦІЯ

Спеціалізована база даних з генної експресії у матково-плацентарному комплексі та крові матері й дитини та плацентарний транскриптом упродовж фізіологічного перебігу вагітності

03.00.03

молекулярна біологія

Подається на здобуття наукового ступеня кандидата біологічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело. \_\_\_\_\_ Лихенко О.К.

Науковий керівник:

**Оболенська Марія Юріївна,**

доктор біологічних наук,

професор

Київ – 2025

## АНОТАЦІЯ

**Лихенко О.К.** Спеціалізована база даних з генної експресії у матково-плацентарному комплексі та крові матері й дитини та плацентарний транскриптом упродовж фізіологічного перебігу вагітності. — Кваліфікаційна наукова праця на правах рукопису.

*Дисертація на здобуття наукового ступеня кандидата біологічних наук за спеціальністю 03.00.03 (молекулярна біологія). — Інститут молекулярної біології і генетики НАН України, Київ, 2025.*

Роботу присвячено створенню спеціалізованої бази даних з генної експресії у зразках матково-плацентарного комплексу та крові матері й плода та інтегративному аналізу транскриптому плаценти упродовж фізіологічної вагітності.

Плацента — тимчасовий орган, що з першого дня вагітності забезпечує обмін продуктами життєдіяльності та імунний діалог між матір'ю та плодом. Порушення формування чи функції плаценти лежать в основі сукупності ускладнень (викидні, затримка росту плода, преєклампсія тощо), які найбільше впливають на материнсько-дитячу захворюваність і смертність. Незважаючи на численні публікації, етіологія та молекулярний патогенез цих ускладнень залишаються не з'ясованими, що обмежує можливості ранньої діагностики й таргетної терапії.

Сучасні високопродуктивні методи дослідження посприяли накопиченню великих масивів різних за походженням даних у відкритих репозиторіях GEO та ArrayExpress. Для проведення аналізу генної експресії в плаценті впродовж фізіологічної вагітності вирішено було звернутися до вторинного аналізу, поширеного в наш час в галузі широкомасштабних досліджень. Першочергово

було вирішено створити спеціалізовану базу даних тільки для зразків матково-плацентарного комплексу і крові матері і дитини зі стандартизованими метаданими як необхідної передумови для проведення вторинного аналізу, який передбачає об'єднання однотипних даних. Метадані для всіх зразків матково-плацентарного комплексу були автоматично отримані з баз ArrayExpress та GEO, поповнені через тексти відповідних статей та листування з авторами і стандартизовані за допомогою Medical Subject Headings (MeSH) та Experimental Factor Ontology (EFO). MeSH використовується для кодифікації ключових слів у біологічних наукових статтях, а EFO — це онтологія, яку ArrayExpress використовує для текстового пошуку під час обробки пошукових запитів користувачів. Таким чином було відібрано, стандартизовано і розширено метадані (діагноз, термін гестації, тип зразків біологічного матеріалу, стать плоду) датасетів з генної експресії в матково-плацентарному комплексі та крові матері й плоду і розроблено реляційну базу даних IGEA (integrative gene expression analysis). Саме стандартизованість метаданих уможливорює простий та швидкий відбір датасетів за обраними користувачем характеристиками зразків і цим відрізняє IGEA від інших подібних спеціалізованих баз даних. Початковий код та резервний файл бази даних **IGEA** доступні на GitHub за посиланням <https://github.com/Sashkow/placenta-preeclampsia>, а веб інтерфейс бази даних доступний за посиланням <https://molecularbiology.ink/igea/>. Структура бази даних дозволяє у майбутньому розширювати її шляхом додавання зразків — навіть із новими типами характеристик — без необхідності змінювати саму структуру бази. Зразки інших компонент матково-плацентарного комплексу зі стандартизованими метаданими вже підготовані для вторинного аналізу, який за бажанням може проводити будь-яка дослідницька група.

Дані обраних з бази IGEA датасетів восьми досліджень генної експресії цілісної тканини плаценти першого, другого та третього триместрів було об'єднано й утворено новий інтегрований датасет. Збільшена кількість зразків у датасеті (основна перевага інтегративного аналізу) та використання оновлених анотацій чипів мікромасивів та ретельна перевірка й усунення небіологічних джерел варіативності даних — як в межах, так і між датасетами — позитивно вплинули на статистичну достовірність результатів у порівнянні з початково взятими дослідженнями.

На підставі даних експресії генів, що були відібрані з обраних з бази IGEA датасетів восьми досліджень цілісної тканини плаценти першого, другого та третього триместрів, було утворено новий інтегрований датасет. За допомогою аналізу диференційної експресії в інтегрованому датасеті отримано оригінальні дані, що кількісно характеризують зміни в експресії 253 генів між першим і другим та 489 генів між другим і третім триместрами вагітності. У першому часовому інтервалі виділено кластери, пов'язані з вродженим імунітетом і хемокінами, морфогенами та компонентами позаклітинного матриксу, ангиогенезом, відповідями на іони  $Zn^{2+}/Cu^{2+}$  та з JAK–STAT-сигналінгом. У другому інтервалі визначено кластери, що охоплюють регуляцію відповіді на зовнішній стимул, метаболізм і клітинний цикл, O-глікозилування, тканинний морфогенез і фактори комітування, іонний транспорт та специфічні для вагітності  $\beta$ -глікопротеїни. Перший часовий інтервал характеризується активними імунними і захисними реакціями, диференціюванням клітин і ростом органу, прозапальними процесами і підтриманням взаємної толерантності у ділянці зіткненням плодової частини плаценти з децидуальною тканиною матері. Другий часовий інтервал характеризується суттєвим зменшенням прозапальних процесів, на заміну яким приходять протизапальні процеси, зменшення проліферативної активності й

активності морфогенів і генів комітування та активація метаболічних процесів, характерних для високодиференційованої тканини.

Вперше описано чотири універсальні траєкторії ко-експресії 54 генів, експресія яких змінювалася впродовж обох досліджуваних періодів вагітності («Up-Up», «Up-Down», «Down-Down», «Down-Up») та характерні схеми взаємодії ДЕГів у межах біологічного процесу: взаємне виключення (наприклад, прозапальні хемокіни проти гомеостатичних) та синергія (активація активаторів морфогенезу з одночасним інгібуванням їх інгібіторів у першій половині вагітності або активація субодиноць інгібіну при інгібуванні активіну в другій половині вагітності).

У другому триместрі вперше показано й кількісно охарактеризовано переломний момент в експресії 24 генів (гени «Up-Down» та «Down-Up»), напрямом експресії від першого до другого триместру змінюється на протилежний від другого до третього триместру. Серед таких генів є, наприклад, ген поліфункціонального лептину (**LEP**), експресія якого знижується від першого до другого триместру і підвищується від другого до третього триместру. Лептин є потенційним маркером преєклампсії. Переломний момент в експресії генів збігається в часі з першим проявом клінічних симптомів, що не виключає залежності клінічних симптомів від дисфункції генів, напрямом експресії яких різний у першій і другій половинах вагітності.

Вперше серед диференційно експресованих генів між першим і другим та другим і третім триместрами фізіологічної вагітності виокремлено гени, які асоційовані з преєклампсією та визначено, які відповідні білки є секретованими.

Проведене дослідження відповідає принципам системної біології. Воно є міждисциплінарним, бо поєднує дані молекулярної біології (з баз даних GEO,

ArrayExpress) з обчислювальними методами комп'ютерних наук та математичними статистичними методами. У використаних даних вимірюється експресія не окремих генів, а профіль генної експресії в цілому. Кількісні дані про зміну експресії генів між триместрами поєднуються із загальними даними про залучення цих генів до біологічних процесів та до клітинних компонентів у людському організмі (база Gene Ontology), мережевих міжбілкових взаємодій (база STRING), секретому (база Human Protein Atlas), а також асоціацій із патологічними станами, зокрема преєклампсією (база CTD). Тому інтегративність аналізу проявляється не лише в об'єднанні однорідних датасетів, а й у поєднанні різнорідних баз біологічних даних, що відображають регуляцію, взаємодію та локалізацію генів. Такий підхід дозволяє оцінити функціонування біологічної системи на молекулярному, клітинному та на рівні організму і втілює принципи системної біології, спрямованої на побудову цілісного уявлення про найскладнішу систему людського організму: мати-плацента-плід.

Ключові слова: плацента, транскриптом, гени, інтегративний аналіз, перший триместр, другий триместр, третій триместр, база даних, молекулярна біологія, метаболізм, імунна регуляція, ангіогенез, гени комітування, морфогени, хемокіни.

## SUMMARY

***Lykhenko O.K.* A specialized database of gene expression in the uterine-placental complex, maternal and fetal blood, and the placental transcriptome throughout physiological pregnancy. — Manuscript. A Dissertation for obtaining the degree of Candidate of Biological Sciences, specialty 03.00.03 — Molecular Biology. — Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, Kyiv.**

*The study is devoted to the creation of a specialized database of gene expression in samples of the uteroplacental complex and maternal and fetal blood and integrative analysis of the placenta transcriptome during physiological pregnancy.*

Open gene expression data accumulated in the Gene Expression Omnibus (GEO) and ArrayExpress repositories were used to conduct a secondary analysis of the combined data of the same type, which increases the sample size and statistical significance of the analysis results. The metadata of the gene expression datasets in the uteroplacental complex and maternal and fetal blood using microarray technology were selected, standardized, and expanded. A specialized relational database of IGEA metadata was created for the first time. The database enables quick and straightforward selection of datasets according to the characteristics of the samples selected by the user.

Gene expression data from the IGEA database, comprising eight studies of bulk placenta tissue across the first, second, and third trimesters, were combined to form a fundamentally new integrated dataset. For the first time, changes in the expression of 253 genes between the first and second trimesters and 489 genes between the second and third trimesters of pregnancy have been quantitatively characterized. Active immune and protective reactions, cell differentiation and organ growth, pro-inflammatory processes, and maintenance of mutual tolerance in the area of contact of the fetal part of the placenta with the decidual tissue of the mother characterize the first time interval. The second time interval is characterized by a significant decrease in pro-inflammatory processes, which are replaced by anti-inflammatory processes, a decrease in proliferative activity and the activity of morphogens and commitment genes, and the activation of metabolic processes characteristic of highly differentiated tissue. In the second trimester, a turning point in the expression of numerous genes has been shown and quantitatively characterized for the first time; the direction of gene expression from the first to the second trimester changes to the opposite direction from the second to the third trimester.

Among such genes is, for example, the polyfunctional leptin gene (LEP), the expression of which decreases from the first to the second trimester and increases from the second to the third trimester. Leptin is a potential marker of preeclampsia, and its protein level is elevated before the onset of clinical symptoms. The turning point in gene expression coincides in time with the first manifestation of clinical symptoms, which does not exclude the dependence of clinical symptoms on the dysfunction of genes, the direction of expression of which is different in the first and second halves of pregnancy.

For the first time, four universal trajectories of gene co-expression (“Up-Up”, “Up-Down”, “Down-Down”, “Down-Up”) and three characteristic patterns of DEG interaction within a biological process have been described: mutual exclusion (for example, pro-inflammatory chemokines versus homeostatic ones), synergy (activation of morphogenesis activators with simultaneous inhibition of their antagonists, and activation of inhibins leads to specific inhibition of activins and suppression of morphogenetic processes).

Keywords: placenta, transcriptome, genes, first trimester, second trimester, third trimester, integrative analysis, database, molecular biology, metabolism, immune regulation, angiogenesis, commit genes, morphogens, chemokines

#### **Список публікацій здобувача:**

1. **Lykhenko O**, Frolova A, Obolenskaya M. Designing the database for microarray experiments metadata. In: 2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF). IEEE; 2017. doi:10.1109/YSF.2017.8126658.
2. **Lykhenko O**, Frolova A, Obolenskaya M. Creation of gene expression database on preeclampsia-affected human placenta. *Biopolym Cell*. 2017;33(6):442-52. doi:10.7124/bc.000967.

3. **Lykhenko O**, Frolova A, Obolenskaya M. Consecutive integration of available microarray data for analysis of differential gene expression in human placenta. *Biotechnol Acta*. 2021;14(1):38-45. doi:10.15407/biotech14.01.38.
4. **Lykhenko O**, Frolova A, Obolenskaya M. Changes in the human placental transcriptome during the physiological course of pregnancy. *Biopolym Cell*. 2021;37(1):73-82. doi:10.7124/bc.000A4D.
5. **Lykhenko O**, Obolenskaya MY. The knowledge of chromosomal sex is important for large-scale analysis of gene expression. *Dopov Nac Akad Nauk Ukr*. 2021;(1):100-9. doi:10.15407/dopovidi2021.01.100.
6. Frolova A, **Lykhenko O**, Bondarenko V, Obolenskaya M. Step by step to integration of gene expression profiles for preeclampsia-affected human placenta. In: China — Ukraine International Symposium on Innovation and Technology; 2017 Oct 30 – Nov 3; Shandong Academy of Sciences. Abstract book. p. 1-8.
7. Zubenko S, **Lykhenko O**, Frolova A, Obolenskaya M. Comparing alternative pipelines for RNA-seq differential expression analysis between preeclamptic and control placental samples. In: XV Всеукраїнська конференція молодих вчених з міжнародною участю, присвячена 30-річчю Незалежності України; 2021 May 26; Kyiv: IMBIG NAS of Ukraine.
8. **Lykhenko O**, Zhoha V, Lykhenko DK, Frolova A, Martsenyuk O, Obolenska M. How smoking may decrease the risk of preeclampsia development. *Med Clin Chem Sci J*. 2019;21(3[80]):110.
9. Zhalilova Y, **Lykhenko O**, Obolenskaya M. Топ-двадцять диференційно експресованих генів у плаценті людини між другим і першим триместрами фізіологічної вагітності. In: Біотехнологія XXI століття: матеріали XVII Міжнародної науково-практичної конференції; 2023 May 19; Kyiv: KPI ім. Ігоря Сікорського. с. 225-7.

10. Kukuruzza Y, **Lykhenko O**, Obolenskaya M. Алгоритм пошуку тканинно-специфічних генів, продукти яких секретуються у кров, на прикладі плаценти людини. In: Біотехнологія XXI століття: матеріали XVII Міжнародної науково-практичної конференції; 2023 May 19; Kyiv: KPI ім. Ігоря Сікорського. с. 228-32.
11. Zhalilova Y, **Lykhenko O**, Obolenskaya M. Pathway enrichment analysis in the human placenta for the period between the second and first trimesters of physiological pregnancy. In: Тези Міжнародної конференції з нейронаук та наукових читань, присвячених вісцеральній фізіології та патофізіології; 2024 Nov 19-21; Kyiv. Fiziol Zh. 2024;70(5 Suppl):33.
12. Martynenko V, Zhalilova Y, **Lyhenko O**, Obolenskaya M. Top 20 differentially expressed genes across trimesters of normal pregnancy. In: Proceedings of the International Scientific Conference BioGENext; 2024 Sep 17-20; Kyiv, Ukraine. Biopolym Cell. 2024;40(3):210. doi:10.7124/bc.000AEA.

## ЗМІСТ

<b>ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ .....</b>	<b>15</b>
<b>ВСТУП.....</b>	<b>16</b>
<b>РОЗДІЛ 1 ОГЛЯД ЛІТЕРАТУРИ .....</b>	<b>22</b>
<b>1.1. Будова і функції плаценти, її розвиток .....</b>	<b>22</b>
<b>1.1.1. Структура зрілої плаценти людини. ....</b>	<b>23</b>
<b>1.1.2. Формування й розвиток плаценти. ....</b>	<b>25</b>
<b>1.1.3. Порушення розвитку і функціонування плаценти — причина багатьох ускладнень вагітності.....</b>	<b>27</b>
<b>1.2. Вагітність і її періоди .....</b>	<b>28</b>
<b>1.3. Системний підхід до дослідження молекулярних процесів у плаценті впродовж фізіологічної вагітності та її ускладнень .....</b>	<b>29</b>
<b>1.4. Інтегративний аналіз — метод для збільшення розміру вибірки й підвищення статистичної значущості результатів .....</b>	<b>30</b>
<b>1.4.1. Інтегративний аналіз даних генної експресії, отриманих за допомогою технології мікромасивів.....</b>	<b>31</b>
<b>1.5. Огляд методів встановлення функціонально збагачених груп генів ..</b>	<b>34</b>
<b>РОЗДІЛ 2 МАТЕРІАЛИ І МЕТОДИ .....</b>	<b>36</b>
<b>2.1. Збір первинних даних .....</b>	<b>36</b>
<b>2.2. Створення спеціалізованої бази даних із генної експресії в плаценті та компонентах матково-плацентарного комплексу.....</b>	<b>37</b>

<b>2.3. Обробка первинних даних генної експресії, встановлення відповідності між пробкетами та генами і визначення показника експресії для кожного гена.....</b>	<b>37</b>
<b>2.3.1. Перевірка якості обробки даних. ....</b>	<b>38</b>
<b>2.4. Нормалізація й інтеграція даних експресії генів в однотипних датасетах .....</b>	<b>39</b>
<b>2.5. Аналіз інтегрованих даних .....</b>	<b>39</b>
<b>2.5.1. Визначення диференційно експресованих генів.....</b>	<b>39</b>
<b>2.5.2. Пошук функціонально збагачених груп диференційно експресованих генів.....</b>	<b>40</b>
<b>2.5.2.1. Оптимізація побудови графа. ....</b>	<b>40</b>
<b>РОЗДІЛ 3 РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ .....</b>	<b>42</b>
<b>3.1. Структура бази даних IGEA.....</b>	<b>42</b>
<b>3.2. Класифікація зразків плаценти й матково-плацентарного комплексу за ключовими метаданими .....</b>	<b>43</b>
<b>3.3. Загальна характеристика генної експресії між першим і другим та другим і третім триместрами нормальної вагітності .....</b>	<b>46</b>
<b>3.4 Оцінювання якості видалення ефекту серійності .....</b>	<b>47</b>
<b>3.5. Визначення і характеристика диференційно експресованих генів між триместрами .....</b>	<b>50</b>
<b>3.6. Кластери диференційно експресованих генів і функціонально збагачені біологічні процеси між першим і другим триместром .....</b>	<b>54</b>
<b>3.6.1. Кластер 1. Імунні процеси. ....</b>	<b>55</b>

3.6.2. Кластер 2. Морфогени, позаклітинний матрикс і ангиогенез.....	57
3.6.3. Кластер 3. Клітинна відповідь на йони цинку й міді.....	62
3.6.4. Кластер 4. Сигналінг через JAK-STAT шлях.....	65
3.7. Кластери диференційно експресованих генів і функціонально збагачені біологічні процеси між другим і третім триместрами .....	67
3.7.1. Кластер 1. Регуляція відповіді на зовнішній стимул.....	67
3.7.2. Кластер 2. Метаболічні процеси і клітинний цикл.....	70
3.7.3. Кластер 3. Процесинг О-гліканів.....	73
3.7.4. Кластер 4. Морфогенез тканин і фактори комітування.....	75
3.7.5. Кластер 5 “Транспорт натрію” і кластер 6 “Трансмембранний транспорт хлориду”.....	80
3.7.6. Кластер 7. Специфічні для вагітності бета-1-глікопротеїни.....	81
3.8. Порівняльний аналіз генної експресії у плаценті першої і другої половини вагітності.....	82
<b>РОЗДІЛ 4 АНАЛІЗ І УЗАГАЛЬНЕННЯ РЕЗУЛЬТАТІВ .....</b>	<b>91</b>
4.1. Створення бази метаданих для зразків матково-плацентарного комплексу в умовах нормального і патологічного перебігу вагітності .....	91
4.2. Порівняльний аналіз генної експресії в плаценті між I і II і між II і III триместрами фізіологічної вагітності.....	92
4.3. Особливості «співробітництва» генів в плаценті та між плацентою і плодом.....	97
4.4. Диференційно експресовані гени, асоційовані з преєклампсією.....	98

<b>ВИСНОВКИ.....</b>	<b>102</b>
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....</b>	<b>105</b>
<b>ДОДАТОК А СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ, В ЯКИХ ОПУБЛІКОВАНІ НАУКОВІ РЕЗУЛЬТАТИ ДИСЕРТАЦІЇ .....</b>	<b>124</b>
<b>ДОДАТОК Б РОЗРОБКА СТРУКТУРИ БАЗИ ДАНИХ.....</b>	<b>127</b>
<b>ДОДАТОК В ДИФЕРЕНЦІЙНО ЕКСПРЕСОВАНІ ГЕНИ МІЖ І-ІІ ТРИМЕСТРАМИ ЗА КЛАСТЕРАМИ .....</b>	<b>132</b>
<b>ДОДАТОК Г ДИФЕРЕНЦІЙНО ЕКСПРЕСОВАНІ ГЕНИ МІЖ ІІ-ІІІ ТРИМЕСТРАМИ ЗА КЛАСТЕРАМИ .....</b>	<b>139</b>
<b>ДОДАТОК Д АСОЦІЙОВАНІ З ПРЕЕКЛАМПСІЄЮ ГЕНИ СЕРЕД ДИФЕРЕНЦІЙНО ЕКСПРЕСОВАНИХ ГЕНІВ МІЖ ПЕРШИМ І ДРУГИМ ТА ДРУГИМ І ТРЕТІМ ТРИМЕСТРАМИ ВАГІТНОСТІ ..</b>	<b>151</b>
<b>ДОДАТОК Е ВЗЯТІ ДЛЯ СТВОРЕННЯ СПЕЦІАЛІЗОВАНОЇ БАЗИ ДАНИХ ІДЕНТИФІКАТОРИ ARRAYEXPRESS .....</b>	<b>153</b>

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

1_2	— порівняння даних між першим і другим триместром вагітності
2_3	— порівняння даних між другим і третім триместром вагітності
ArrayExpress	— база даних функціональної геноміки
CDF	— файли анотації генів для платформи Affymetrix
CEL	— файли первинних даних мікромасивів Affymetrix
FC	— fold change, кратність зміни експресії гена
FDR	— false discovery rate (adjusted p-value), частка хибних відкриттів (скориговане p-значення)
GEO	— Gene Expression Omnibus, база даних
GO	— Gene Ontology, база даних
GSE	— GEO Series, ідентифікатор датасету в базі даних GEO
LogFC	— logarithm fold change, логарифм кратності зміни (FC) за основою 2
PCA	— principal component analysis, аналіз головних компонент
SQL	— structured query language, структурована мова запитів до баз даних
STRING	— база даних білок-білкових взаємодій і її веб інтерфейс
Датасет	— структурований набір даних, який використовується для аналізу, машинного навчання або побудови гіпотез
ДЕГ	— диференційно експресований ген
Ефект серійності	— в аналізі даних ефект серії стосується систематичних, небіологічних відмінностей між серіями зразків, які можуть спотворити результати досліджень
ПЗКМ	— позаклітинний матрикс (англ. ECM, extracellular matrix)
ФЗБП	— функціонально збагачений біологічний процес (англ. functionally enriched biological process)

## ВСТУП

**Актуальність теми.** Плацента — тимчасовий орган, що з першого дня вагітності забезпечує обмін продуктами життєдіяльності та імунний діалог між матір'ю та плодом. Порушення формування чи функції плаценти лежать в основі «великого акушерського синдрому» — сукупності ускладнень (викидні, затримка росту плода, прееклампсія тощо), які найбільше впливають на материнсько-дитячу захворюваність і смертність. Незважаючи на численні публікації, етіологія та молекулярний патогенез цих ускладнень залишаються не з'ясованими, що обмежує можливості ранньої діагностики й таргетної терапії [1, 2].

Сучасні високопродуктивні методи дослідження сприяли накопиченню великих масивів транскриптомних даних плаценти у відкритих репозиторіях GEO та ArrayExpress. Однак більшість досліджень охоплює небагато зразків і різняться за дизайном та анотацією, тож окремо не забезпечують належної статистичної сили. Інтегративний аналіз первинних мікроарей-даних на відміну від метааналізу дає змогу подолати цей бар'єр. Після стандартизації метаданих, нормалізації та корекції ефекту серійності об'єднання даних генної експресії з різних, але однотипних зразків плаценти, отримуємо єдиний якісний датасет, який дозволяє виявити гени, що статистично достовірно диференційно експресуються між триместрами, та зрозуміти, як саме плацентарний транскриптом перебудовується протягом нормальної гестації. Таке знання формує міцний молекулярний ґрунт для подальших досліджень механізмів патологічних вагітностей і розробки біомаркерів для раннього прогнозування ускладнень і таргетної терапії.

**Об'єктом дослідження** є первинні дані зразків матково-плацентарного комплексу та крові матері й дитини, які знаходяться у відкритому доступі й охарактеризовані за генною експресією. **Предметом дослідження** є зміни в

інтегрованому білок-кодуючому **транскриптомі плаценти між триместрами** впродовж фізіологічного перебігу вагітності.

**Зв'язок роботи з науковими програмами, планами, темами.** Роботу виконано відповідно до тематики групи системної біології відділу ензимології білкового синтезу, що затверджена наказом від 26.11.2020р.№4, а також у рамках українсько-австрійського гранту 2023 - 2024 рр. № 0123U104257.

**Мета роботи** — створити спеціалізовану базу метаданих для зразків матково-плацентарного комплексу та крові матері й дитини, які знаходяться у відкритому доступі й охарактеризовані за генною експресією, і визначити зміни у плацентарному транскриптомі впродовж нормального перебігу вагітності у відібраних і об'єднаних зразках методом інтегративного аналізу.

Для досягнення поставленої мети було сформульовано **завдання дослідження:**

1. Створити вторинну спеціалізовану базу метаданих для зразків матково-плацентарного комплексу та крові матері й дитини, використовуючи бази GEO і ArrayExpress, які містять як метадані, так і дані генної експресії.
2. Відібрати з бази IGEA однотипні за метаданими зразки і отримати з баз GEO і ArrayExpress відповідні дані генної експресії за ідентифікаційними номерами досліджень.
3. Перетворити первинні дані експресії (дані про інтенсивність випромінювання на пробах мікромасивів) у логарифмовані та нормалізовані матриці експресії, видалити ефект серійності та провести інтеграцію даних генної експресії.
4. Верифікувати видалення ефекту серійності методом головних компонент.

5. Визначити диференційно експресовані гени (ДЕГи) між I і II та II і III триместрами вагітності.
6. За даними з бази STRING побудувати відповідні ДЕГам мережі міжбілкових взаємодій і асоціацій та кластеризувати вузли (білки) мереж.
7. Визначити біологічні процеси, на які збагачені гени у кожному з кластерів.
8. Охарактеризувати функціонально збагачені на біологічні процеси гени і виявити закономірності змін в експресії генів у межах цих процесів.

У дослідженні застосовано комплекс **методів** біоінформатичного аналізу: *збір і стандартизація даних* (бази даних GEO, ArrayExpress; категоризація метаданих за онтологіями MeSH і EFO); програмні засоби Python/Django/Postgres для розробки веб-бази IGEA; *обробка даних експресії* (RMA в R пакеті affy для .CEL-файлів Affymetrix; фон-корекція,  $\log_2$ -трансформація й квантильна нормалізація в limma для Illumina); *інтеграція даних* (емпіричний метод Байєса реалізований у методі ComBat з R пакета sva); *контроль якості* первинної обробки даних (PCA, t-SNE); визначення *диференційно експресованих генів* (t-тест, лінійні моделі, обчислення логарифму кратності зміни  $\log_{FC}$  з FDR-корекцією); *кластерний аналіз* (ієрархічна кластеризація «fast-greedy» на графі білок-білкових взаємодій STRING); *визначення функціонально збагачених груп генів* (гіпергеометричний тест, за даними з баз STRING і Gene Ontology); побудова теплокарт близькості біологічних процесів (індекс Джаккарда, R пакет rheatmap).

**Наукова новизна.** Створено й опубліковано на веб-сайті molecularbiology.ink базу IGEA, яка надає стандартизовані метадані для зразків матково-плацентарного комплексу та крові матері й плоду і надає можливість для проведення вторинного аналізу генної експресії. За допомогою створеного

інструменту виконано інтегративне об'єднання даних мікромасивів за 20 останніх років, отримано новий узагальнений датасет і вперше надана кількісна характеристика рівня експресії достовірних ДЕГів та встановлено біологічні процеси, на які функціонально збагачені ці ДЕГи, для інтервалів перший-другий (1\_2) та другий-третій триместри (2\_3); показано, що на межі першої і другої половини вагітності траєкторії експресії численних генів «переламуються» (зокрема маркерний LEP), що збігається з часом появи клінічних симптомів преєклампсії. Описано чотири універсальні траєкторії ко-експресії генів («Up-Up», «Up-Down», «Down-Down», «Down-Up») та три характерні схеми взаємодії ДЕГів у межах біологічного процесу: взаємне виключення (наприклад, прозапальні хемокіни проти гомеостатичних), синергія (активація активаторів морфогенезу з одночасним інгібуванням їх антагоністів та активація інгібінів призводить до специфічного інгібування активінів і пригнічення морфогенетичних процесів). Разом ці спостереження уточнюють молекулярний сценарій нормального розвитку плаценти.

**Особистий внесок автора** в роботу був визначальним на етапах відбору датасетів, стандартизації метаданих, створення бази даних і веб-сайту з доступом до неї, біоінформатичного аналізу даних генної експресії, і суттєвим — на етапі біологічної інтерпретації диференційної експресії генів. Автор вдячний науковому керівникові Оболенській М.Ю за супровід на всіх етапах виконання роботи та, зокрема: за допомогу в підборі вдалих термінів з онтологій EFO і MESH при створенні бази метаданих IGEA (публікація «Creation of gene expression database on preeclampsia-affected human placenta»); за додаткову верифікацію коректності роботи пайплайну інтегративного аналізу при його розробці шляхом оцінки, чи мають ті чи інші результати, які видає пайплайн, біологічний зміст (публікація

«Consecutive integration of available microarray data for analysis of differential gene expression in human placenta»); за значну допомогу в біологічній інтерпретації встановленої диференційної експресії генів та у виявленні цікавих закономірностей змін генної експресії в межах різних біологічних процесів (публікації «Changes in the human placental transcriptome during the physiological course of pregnancy» та «The knowledge of chromosomal sex is important for large-scale analysis of gene expression») та за численні правки та рекомендації при підготовці публікацій до друку. Автор висловлює подяку м.н.с. Фроловій А. О. за перевірку коректності застосованих в усіх згаданих публікаціях біоінформатичних методів. Автор також вдячний С. А. Фаундс та іншим авторам датасетів, які в ході особистого листування поповнили наявні дані додатковими неопублікованими метаданими.

**Апробація матеріалів дисертації** проводилася на наукових заходах: 17–20 жовтня 2017 р. на 2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF-2017), м. Львів, Україна EPS Young Minds; 30 жовтня – 3 листопада 2017 р. на China–Ukraine International Symposium on Innovation and Technology, Shandong Academy of Sciences, Циндао, Китай; 26 травня 2021 р. на XV Всеукраїнській конференції молодих вчених з міжнародною участю, Інститут молекулярної біології і генетики НАН України, Київ; 19 травня 2023 р. на «Біотехнологія XXI століття»: XVII Міжнародній науково-практичній конференції, КПІ ім. Ігоря Сікорського, Київ; 19–21 листопада 2024 р. на Міжнародній конференції з нейронаук та наукових читаннях, Київ; 17–20 вересня 2024 р. на International Scientific Conference BioGENext, Київ, Україна. Ці виступи були публічною апробацією підходів та результатів дослідження серед фахової спільноти.

**Публікації.** За темою дисертації опубліковано 5 наукових публікацій: дві статті індексовано у міжнародній наукометричній базі даних Scopus (Q4), одну статтю представлено у виданнях, що входять до інших міжнародних наукометричних баз даних, дві статті опубліковано у наукових фахових виданнях України; 7 тез доповідей на міжнародних та всеукраїнських науково-практичних конференціях, симпозіумах і конгресах.

**Структура та обсяг роботи.** Дисертація складається з анотації, переліку публікацій здобувача, вступу, огляду літератури, матеріалів і методів дослідження, результатів інтегративного біоінформатичного аналізу, розділу з обговоренням та узагальненням отриманих результатів, висновків, списку використаних джерел і додатків. Роботу викладено на 153 сторінках тексту; вона містить 32 рисунків, 3 таблиці та 6 додатків.

## РОЗДІЛ 1

### ОГЛЯД ЛІТЕРАТУРИ

#### 1.1. Будова і функції плаценти, її розвиток

Плацента — це орган, що об'єднує мати і плід у єдину систему мати-плацента-плід і формує гемоплацентарний бар'єр. Плацента забезпечує живлення й постачання кисню до плоду і виведення оксиду вуглецю та продуктів обміну речовин плоду в кров матері, вона забезпечує ріст і розвиток плоду, зумовлює зміни в організмі матері, необхідні для підтримання вагітності, і перешкоджає відторгненню напівалогенного плоду. Плацента є регулятором внутрішньоутробного середовища для плоду й частково виконує функції майже всіх систем людського організму — печінки (обмін речовин, детоксикація), ендокринної системи (синтез гормонів), легенів (обмін кисню й оксиду вуглецю), імунної системи (захист плоду від імунного конфлікту з організмом матері і від інфекцій) і нирок (фільтрувальна функція й сольовий обмін) [3]. Плацента бере активну й пасивну участь в обміні речовин матері й плоду. Активна участь полягає в продукції регуляторних речовин, які представлені пептидними та стероїдними гормонами, факторами росту, цитокінами тощо. Наприклад, з перших днів запліднення в кров матері потрапляє хоріонічний гонадотропін (human chorionic gonadotropin, hCG), який продукується трофобластом і за яким визначають стан вагітності [4]. Плацентарний лактоген (placental lactogen, PL) поступає в кровоток матері та у незначній кількості у кровоток плоду, де він активує інсуліновий ростовий фактор (IGF) і його ізоформи [5]. Плацента активно транспортує метаболіти. Частина метаболітів, включно з киснем і оксидом вуглецю, потрапляють у кровообіг відповідно плоду й матері завдяки пасивній дифузії. Плацента бере активну участь у програмуванні плоду, тобто за певних обставин

опосередковує такі зміни в розвитку плоду, які можуть зумовлювати ті чи інші захворювання в подальшому житті людини [6].

Порушення у функціонуванні плаценти негативно впливають як на організм матері, так і на розвиток ембріона й плоду, стають причиною ускладнень вагітності — викиднів, завмерлої вагітності, затримки внутрішньоматкового росту, прееклампсії [2]. Деякі дисфункції плаценти є критичними для життя матері й дитини, наприклад, у разі прееклампсії.

**1.1.1. Структура зрілої плаценти людини.** Зріла плацента людини має дві частини: материнську, прилеглу до міометрію децидуальну оболонку (*decidua basalis*), і плодову, утворену хоріоном (похідним від трофобласта — поверхневого шару бластоцисти), сполучною тканиною, кровеносними судинами й амніоном (рис. 1.1) [7]. Плацента не містить нервів чи лімфатичних судин. Через її кровеносні судини відбувається кровообіг між плацентою і плодом. На поверхні ворсин хоріону може знаходитися фіброїд Лангханса, неклітинна фібриноподібна маса із залишків розпаду епітелію ворсин та згортання плазми материнської крові. Фіброїд Лангханса з'являється на поверхні ворсин у другій половині вагітності.

Амніотична оболонка складається із шару епітелію та шару щільної волокнистої сполучної тканини. Між амніотичною оболонкою та хоріоном є шар слизової сполучної тканини.

Хоріон включає хоріальну пластину та ворсини, що виходять із хоріальної пластини і спрямовані вбік *decidua basalis*. Стволові ворсини, які виходять безпосередньо із хоріальної пластини, розгалужуються на ворсинки другого та третього порядку. Стволова ворсина разом з усіма її розгалуженнями утворює котиледон. Виокремлюють якірні ворсинки, які доходять базальної частини ендометрію та зафіксовані в ньому. Інші ворсини занурені в лакуни, заповнені материнською кров'ю.

Хоріон, подібно до амніона, складається зі сполучнотканинної та епітеліальної частини. Сполучна тканина ворсин та хоріальної пластини містить фібробласти й макрофаги, колагенові волокна (трапляються порівняно рідко), розгалужену мережу пупкових судин, у тому числі капіляри, що прилягають до епітелію ворсин. Епітеліальна частина хоріона за походженням є трофобластом і має шар клітин цитотрофобласту, які розташовані на базальній мембрані ворсин і зберігають мітотичну активність, та поверхневий багатоядерний шар синцитіотрофобласту, який виникає під час злиття клітин цитотрофобласту. Синцитіотрофобласт вистилає усю поверхню ворсин хоріону, окрім місць контакту якірних ворсин із шаром ендометрію. Описані структури плодової частини плаценти утворюють гемоплацентарний бар'єр, який відокремлює кров матері від крові плоду. Материнська частина плаценти не задіяна у формуванні гемоплацентарного бар'єру.

Материнська частина плаценти представлена *decidua basalis*, яка утворює лакуни, заповнені материнською кров'ю, та сполучнотканинні септи — перегородки між лакунами. Всередині перегородок проходять судини матері, що відкриваються до лакун. Частина *decidua basalis*, базальна пластинка, формується з нижніх шарів ендометрію і включає сполучну тканину і скупчення децидуальних клітин і міофіброblastів. Якірні ворсинки хоріона закріплюються на базальній пластинці. Завдяки клітинам міофіброblastу *decidua basalis* відбувається регенерація сполучної тканини після пологів. Епітелій відновлюється використовуючи ділянки ендометрію, де залишилися денця маткових залоз. Базальна пластинка інфільтрована децидуальними імунними клітинами, до яких належать природні клітини — вбивці, Т-клітини, макрофаги, які мають кістковомозкове походження [8]. Ці клітини виробляють гормон релаксин, що слугує для підготовки материнських тканин і органів до пологів, та мають

макрофагову активність, що перешкоджає надмірному росту ворсин хоріону під час вагітності. Концентрація релаксину суттєво зростає під час пологів, сприяючи відторгненню плаценти.

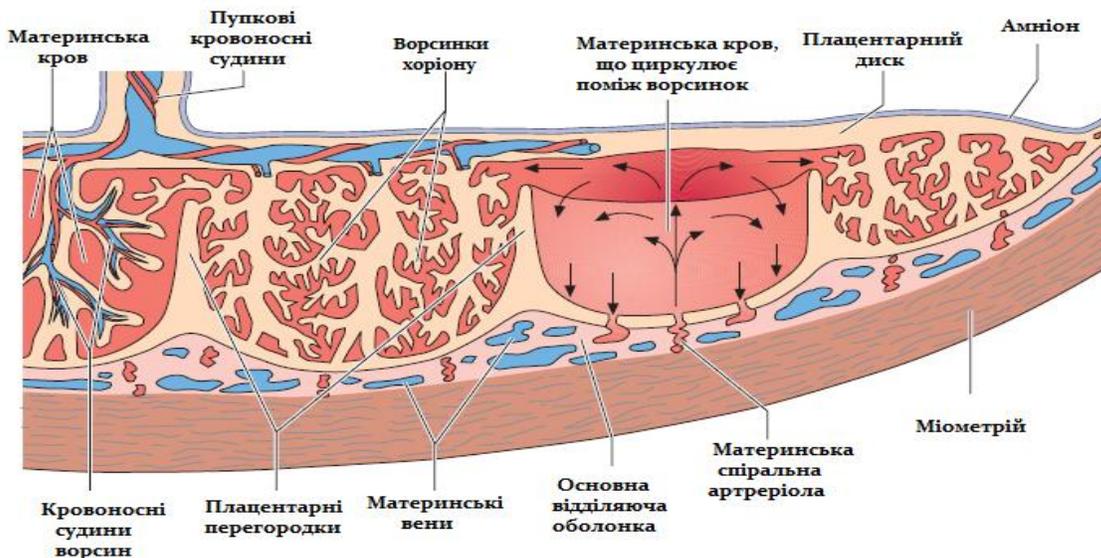


Рис. 1.1. Схематичне зображення плаценти людини за [7].

**1.1.2. Формування й розвиток плаценти.** Формування плаценти є складним процесом, що потребує координованого регулювання інвазії трофобласта в кровоносні та лімфатичні судини матки, проліферації й диференціації клітин трофобласта й одночасних змін у децидуальній оболонці матері [2].

На ранніх етапах вагітності в центральній ділянці плодової частини плаценти формуються тонкі вирости, які згодом починають інтенсивно галузитись і утворювати ворсини хоріону другого і третього порядку. Ворсини зрілої плаценти занурені в лакуну материнської частини плаценти й постійно омиваються кров'ю, що забезпечує належний транспорт речовин між матір'ю і плацентою [9]. У людини на першому триместрі вагітності, ворсинки плаценти складаються з мезенхімних клітин, що оточені шаром клітин цитотрофобласту (внутрішній, клітинний шар трофобласту). Клітини цитотрофобласту зливаються і формують синцитіотрофобласт, зовнішній багатоядерний клітинний шар

трофобласту. Синцитіотрофобласт — це спеціалізований епітелій, який контактує з материнською кров'ю і через нього здійснюється двосторонній обмін газами й поживними речовинами від матері до плоду та продуктами життєдіяльності плоду до матері [10]. Близько 17-го дня після запліднення з'являються ворсини третього порядку та первинні кровоносні судини, які походять із гемангіогенних центрів. Останні розвиваються з позаембріонального мезодерма ще задовго до появи кровоносних судин у плода (32-й день після запліднення). Розгалуження кровоносних судин і утворення капілярів (ангіогенез) продовжується у другому триместрі, а їхнє об'єднання із судинами плоду в єдину систему плацентарно-ембріонального кровообігу відбувається після 4-го тижня [2]. Червоні кров'яні тіลця і клітини Гофбауера наявні в плаценті на 18-й день після запліднення. Вважають, що перші утворюються внаслідок трансформації ендотеліальних клітин, а клітини Гофбауера зі стромальних клітин-попередників, які схожі на фібробласти, але мають маркери макрофагів/моноцитів [11].

На початку вагітності з ворсинок хоріону походить позаворсинковий цитотрофобласт, утворює колони (тяжі тробласту, які з'єднують сусідні ворсинки і скріплюють їх) і занурюється в децидуальну оболонку, де взаємодіє з клітинами строми, клітинами-вбивцями й макрофагами децидуальної оболонки й регулює імунну відповідь клітин матері на клітини плоду. У позаворсинковому цитотрофобласті виокремлюють ендоваскулярний (endovascular cytotrophoblast, eCTB) й інтерстиціальний цитотрофобласт (interstitial cytotrophoblast, iCTB). Ендоваскулярний цитотрофобласт проникає в спіральні артерії матері і трансформує їх (ремоделює) [12], але до 10 тижня надходження материнської крові до плаценти заблоковане пробкою з позаворсинкового тробласту. Інтерстиціальний цитотрофобласт проникає в струму децидуальної оболонки й у децидуальні залози, лімфатичні судини і вени. Таким чином забезпечується

гістіотрофне живлення плоду завдяки залозам, багатим на поживні речовини, фактори росту й цитокіни, і відтік рідини від плаценти через лімфатичні судини і вени матері, коли кровообмін мати-плацента іще відсутній. У цей період плацента і плід розвиваються в гіпоксичному середовищі і відбувається адаптація материнського організму до клітин плоду [13, 2]. Морфологічні зміни в плаценті є відображенням мільйонів років адаптації до змін у концентрації кисню в земній атмосфері [14]. До кінця першого триместру структура плаценти формується остаточно, що збігається з постачанням материнської крові до плаценти й супроводжується окислювальним стресом через супутнє підвищення концентрації кисню. Ближче до кінця вагітності поверхня плаценти досягає 11 — 14 м<sup>2</sup>, шар епітелію хоріона стає тоншим, зростає щільність ядер синцитіотрофобласту, капіляри розростаються та ближче прилягають до поверхні ворсин, що сприяє найінтенсивнішому обміну між матір'ю і плодом через плаценту.

**1.1.3. Порушення розвитку і функціонування плаценти — причина багатьох ускладнень вагітності.** З порушеннями у функціонуванні плаценти пов'язують низку ускладнень вагітності. Преєклампсія — поширене у світі захворювання (5–10 % вагітностей) з найвищим рівнем смертності матері і плоду та проблемами зі здоров'ям упродовж подальшого життя обох. Етіологія преєклампсії досі невідома, патогенез не до кінця зрозумілий, за браком ранньої діагностики і специфічного лікування передчасне розродження є єдиним заходом, який вирішує проблему. Однак, воно є ризикованим для недоношеної дитини й обтяжливим для системи охорони здоров'я у зв'язку з доглядом за такими немовлятами. Затримкою росту плоду (Intrauterine growth restriction, IUGR) означають випадки, коли маса плоду нижча за 10-ту перцентиль для його строку гестації, тобто коли його маса нижча за мінімальне значення маси у 10 % плодів, яка спостерігається на даному терміні гестації [23]. Преєклампсія часто

супроводжується затримкою росту плоду. Втрата вагітності до 20-го тижня класифікується як спонтанний аборт, а в більш пізні строки — як завмерла вагітність.

Є багато причин, через які відбуваються такі ускладнення вагітності. Серед них першочергову роль відіграють порушення у функціонуванні плаценти, як, наприклад, недостатня інвазія трофобласту й неповне ремоделювання маткових артерій у разі прееклампсії й затримки росту плоду [24]. Ці ускладнення майже неможливо рано діагностувати. Значною мірою це пов'язано з тим, що на наш час ще недостатньо системних відомостей щодо функціонування плаценти в триєдиній системі мати-плацента-плід упродовж фізіологічної й патологічної вагітності.

## 1.2. Вагітність і її періоди

Вагітність — це динамічний процес, який відбувається в організмі матері, плоду й у плаценті з початку запліднення до пологів. Строк вагітності визначають двома способами. У клініці прийнято визначати строк гестації (gestational age, GA) — від першого дня останнього менструального циклу в тижнях, місяцях і триместрах. Інший спосіб — це вік плоду, який визначається з моменту запліднення, і він на 2 тижні коротший за строк гестації. Цю розбіжність необхідно враховувати, адже за 2 тижні численні зміни відбуваються в триєдиній системі мати-плацента-плід [25]. За фізіологічних умов вагітність у людини триває 40 — 42 тижні або три триместри.

Упродовж розвитку плоду виділяють вікна вразливості (windows of susceptibility, WOS) — це часові періоди, які збігаються із періодом розвитку певного органу, коли експресія генів найбільш вразлива до дії різних за походженням несприятливих факторів. Епігенетичні зміни, які можуть виникнути

в період підвищеної вразливості, зумовлюють ризик виникнення дисфункції/хвороби в подальшому житті людини й навіть передачі цих змін до наступного/их поколінь. Ступінь вразливості залежить від генетичних факторів і від статі плоду. Концепція про вікна вразливості відома як гіпотеза Баркера (the Barker Hypothesis) або Розвиток і Витоки Здоров'я і Хвороб (Developmental Origins of Health and Disease, DOHaD) [26].

### **1.3. Системний підхід до дослідження молекулярних процесів у плаценті впродовж фізіологічної вагітності та її ускладнень**

Суттєвий поступ у розвиткові високопродуктивних і комп'ютерних технологій відкрив шлях до впровадження системних підходів під час дослідження молекулярних процесів у плаценті людини за умови фізіологічного й патологічного перебігу вагітності. На сьогодні накопичені дані щодо епігеному, асоціації між фенотиповими ознаками і структурою геному (genome-wide association studies, GWAS), експресії генів, які включають особливості плацентарного транскриптому, протеому й метаболізму. Результати цих досліджень, а також і об'єднані дані з різних омів (multi omics) наведені в декількох систематичних оглядах [27, 28, 29, 30, 31, 32]. Переважна кількість даних стосується профілів генної експресії, оскільки мікроарей технології з олігонуклеотидними послідовностями почали використовувати раніше за всі інші омівки. Незважаючи на інтенсивний розвиток секвенування нового покоління, дослідження генної експресії в суцільній тканині плаценти й компонентах матково-плацентарного комплексу за допомогою мікроарей технологій складають поки що більшу частку серед усіх досліджень. Наприклад, база ArrayExpress (BioSamples) містить на грудень 2023 року 70 досліджень (3046 зразків) генної експресії в плаценті людини, отримані методом секвенування РНК (в базі даних позначені як rna-seq of coding rna), й 287 досліджень (25730 зразків), отриманих за

використання мікроарей технологій (transcription profiling by array). Отримати актуальні дані можна за посиланням <https://www.ebi.ac.uk/biostudies/arrayexpress/studies?query=placenta>

Результати досліджень за використання високопродуктивних методів мають один спільний недолік — найчастіше вони здійснюються на невеликій кількості зразків через вартість дослідження, що зумовлює недостатню надійність висновків. Дослідження генної експресії в плаценті людини мають такий самий недолік. Крім цього, через етичні передумови для дослідження найчастіше використовують зразки плаценти після пологів. Однак у плаценті відбуваються зміни впродовж усього періоду гестації. Зміни на початку гестації можуть зумовлювати розвиток патологічних процесів і мати значення для ранньої діагностики, а процес пологів вносить свої особливості в експресію генів. Бажано прослідкувати зміни в експресії генів упродовж усього періоду гестації в усіх компонентах матково-плацентарного комплексу всіма доступними методами високопродуктивних технологій і в контексті зі змінами в організмі матері й плоду в рамках широкомасштабних досліджень. Поки що таких досліджень немає, але є поступ у цьому напрямку.

#### **1.4. Інтегративний аналіз — метод для збільшення розміру вибірки й підвищення статистичної значущості результатів**

Як зазначалося вище, узв'язку з високою вартістю високопродуктивних технологій дослідники використовують малі вибірки зразків. Щоби підвищити статистичну достовірність результатів необхідно проводити широкомасштабні дослідження, які коштують дуже дорого, або об'єднувати дані наявних досліджень, додаючи до них ті нові, які постійно з'являються. Першочерговими критеріями для об'єднання є подібність клінічних характеристик, наприклад,

діагноз, вік породіллі, строк вагітності тощо. Об'єднанню підлягають дані, отримані однією технологією на однаковій структурі, наприклад, дані з генної експресії в суцільній тканині плаценти (горизонтальне об'єднання), і дані, отримані на різних рівнях клітинної організації, наприклад, дані з транскриптому і протеому в суцільній тканині плаценти (вертикальне об'єднання, *multi omics*), які віддзеркалюють різні аспекти діяльності клітини. Кожне з цих об'єднань потребує використання досить складного математичного апарату.

Є два способи об'єднання — метааналіз і інтегративний аналіз. Метааналіз — метод, який ґрунтується на об'єднанні *результатів* кількох досліджень, а інтегративний аналіз — на об'єднанні *початкових даних* кількох досліджень із наступним аналізом об'єднаних даних. Інтегративний аналіз має низку переваг порівняно з метааналізом, а саме всі дослідження аналізуються за загальною схемою, результати аналізу мають кількісні характеристики. Інтегративний аналіз є технічно складнішим за метааналіз.

**1.4.1. Інтегративний аналіз даних генної експресії, отриманих за допомогою технології мікромасивів.** Для об'єднання даних із генної експресії також використовують метааналіз [32, 33, 34] і інтегративний аналіз [35]. У випадку метааналізу об'єднують списки диференційно експресованих генів із декількох досліджень. У разі ж проведення інтегративного аналізу особливого значення набуває первинна обробка даних, яка передує їхньому об'єднанню. У мікроарей технології найчастіше використовуються два типи чипів із мікромасивами олігонуклеотидних послідовностей, від фірми Affymetrix і фірми Illumina. Чип від Affymetrix (*gene chip*) містить від 500 000 до ~ 6 500 000 комірок із різними послідовностями або пробами. Кожна комірка містить багато однакових копій олігонуклеотидної послідовності завдовжки 25 п.н. Набір із 8–16 проб, що відповідає одному й тому ж гену, зветься пробсетом [36]. Чипи Illumina (*gene*

beads) представлені кремнієвими намистинками з пришитими до них олігонуклеотидними послідовностями завдовжки 50 нуклеотидів. На кожній намистинці біля 30 копій послідовностей, які слугують як технічні повтори.

Первинна обробка даних із мікрочипів включає аналіз якості даних на чипі, оновлення анотації пробсетів і встановлення взаємно однозначної відповідності між пробсетами й генами. Чипи мікромасивів розробляють із метою забезпечити взаємно однозначну відповідність між генами та пробами. Утім, відповідно до того, як наші знання зростають, виявляється, що деякі проби відповідають декільком генам, що актуалізує повторну анотацію мікромасивів згідно з новими біологічними даними. Для оновлення анотації генів на мікрочипах Affymetrix використовують підхід, який розроблений у проєкті Brainarray (<http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/>) і передбачає такі критерії добору проб: кожна проба має відповідати лише одній ділянці геному; проби, що відповідають різним ділянкам однієї кДНК у напрямку 5'→ 3', групуються в єдиний пробсет; кожен пробсет має містити хоча б три олігонуклеотидні проби й ці проби мають бути впорядковані згідно з їхнім розташуванням у відповідному екзоні. Оновлена в такий спосіб анотація мікрочипів Affymetrix забезпечує взаємно однозначну відповідність між пробсетами й генами. Для чипів Illumina обирається проба, яка має найбільше середнє значення інтенсивності між зразками в датасеті серед інших проб, що відповідають одному й тому ж гену, як це робиться в [37, 38], встановлює взаємно однозначну відповідність між пробсетами та генами з найменшими втратами інформації.

Для того, щоб інтеграція мала сенс, кожен зразок має бути описаним із клінічної та біологічної точки зору, і чим повніший цей опис, тим більше інформації можна дістати. Потрібно знати, за якими характеристиками біологічні

зразки схожі, а за якими відрізняються. Це дає можливість обрати зразки для об'єднання так, щоби вони відрізнялися лише за однією важливою характеристикою, наприклад, за наявністю/відсутністю певної патології, і були максимально близькі за рештою характеристик, що надає можливість з'ясувати, чим відрізняється в такому разі патологічний процес від фізіологічного. Сукупність клінічних та біологічних характеристик називається *метаданими* (не плутати з метааналізом, методом об'єднання результатів кількох досліджень).

Після проведення первинної обробки даних важливим кроком у інтегративному аналізі є нормалізація даних між різними датасетами. Таблиці значень генної експресії датасетів з двох різних досліджень не можна просто приєднати одну до одної, зіставивши відповідні рядки-гени щоби утворити інтегрований датасет. Значення експресії генів в обох датасетах відповідає *відносним* значенням інтенсивності випромінювання комірок мікромасиву. Інтенсивність випромінювання систематично відрізняється між двома датасетами з технічних причин, подібно до світлин одного й того ж самого об'єкта, зроблених за різного освітлення різними фотокамерами. Ці систематичні технічні відмінності звуться *ефектом серійності* (англ. batch effect) та можуть бути усунені коригуванням параметрів статистичного розподілу значень генної експресії для кожного зразка, щоби усі зразки мали однакове математичне сподівання й дисперсію. У разі, коли розподіли не є нормальними, коригуються значення розташування й масштабу (location and scale) розподілів, що є узагальненням математичного сподівання та дисперсії.

Інший підхід до видалення ефекту серійності — це розбити зразки в кожному датасеті на підгрупи відповідно до біологічних метаданих та скоригувати коефіцієнти зсуву та масштабу в підгрупах. Наприклад, припустимо ми маємо два датасети, що становлять серії A та B. Серії A і B містять зразки в нормі ( $A_N$ ,  $B_N$ )

та за умов патології (AD, BD) відповідно. Ідея в тому, щоби скоригувати, розподіли генної експресії в підгрупах AN і BN та AD і BD відповідно, замість того щоби робити корекцію для серій A і B загалом. Цей принцип лежить в основі функції ComBat з R пакету sva, що є реалізацією емпіричного методу Байєса [39, 40], який ми обрали для видалення ефекту серійності.

Ефект серійності майже завжди виникає при поєднанні датасетів з кількох різних досліджень, проте його іноді можна спостерігати й в межах одного датасета, коли існує якась технічна причина систематичної різниці в експресії зразків, наприклад, коли гібридизацію на мікромасивах проводили у кілька серій в різні дні.

Кінцевою метою дослідження генної експресії в різних за метаданими зразках є визначення диференційно експресованих генів (ДЕГів) і їхній аналіз для створення гіпотез щодо особливостей процесів, які порівнюються.

## **1.5. Огляд методів встановлення функціонально збагачених груп генів**

Аналіз диференційної експресії може давати велику кількість ДЕГів, що може ускладнити біологічну інтерпретацію результатів. Відомо, що існує когнітивна похибка людини, що проводить дослідження, надавати більше уваги тим генам і їх властивостям, які підтверджують її попередні припущення. Існує формальний автоматизований спосіб отримати біологічний опис ДЕГів: встановлення функціонально збагачених груп генів (functional enrichment analysis). Цей метод використовує список експериментально отриманих ДЕГів, списки функціонально-специфічних груп генів (наприклад, список генів імунної системи, апоптозу, метаболізму ліпідів) та визначає за допомогою статистичного тесту (у нашому випадку — гіпергеометричного тесту) чи є співпадінням попадання диференційно експресованих генів у кожен зі списків функціонально-специфічних груп генів.

Гени, що статистично достовірно потрапляють до функціонально-специфічної групи генів, називаються функціонально збагаченими групами генів.

## **1.6. Обґрунтування необхідності створення спеціалізованої бази даних**

На жаль, найбільш відомі відкриті бази GEO і ArrayExpress не завжди містять достатню кількість метаданих. На те є кілька причин. Головна з них — бази не задумувалися з наміром у подальшому інтегрувати дані. Також, оскільки різні дослідницькі групи завантажували дані незалежно одна від одної за браком загальних вимог щодо термінів (найменування), метадані виявилися нестандартизованими і, відповідно, не готовими до автоматичної обробки через інформаційні запити. Наприклад, строк гестації може бути вказаний як «вік», «вік плоду», «вік плоду (в тижнях)», тож строк гестації для всіх зразків плаценти в базі даних неможливо дістати автоматично простим пошуковим запитом. Вибір значущих характеристик був також відданий на поталу тим, хто завантажував дані, тож обрані ними характеристики не завжди збігаються із необхідними для інтеграції даних. Отже, для інтегративного аналізу бажано збирати додаткові метадані та зводити їх і наявні метадані до спільного стандарту. Безперечною перевагою перед загальними базами даних мають спеціалізовані бази даних із власне встановленими критеріями відбору й характеристикою проб.

Мати-плацента-плід є найскладнішою за всі системи в організмі, тож її дослідження передбачає накопичення й оперування величезним обсягом даних, що зумовлює нагальну необхідність створення спеціалізованої бази даних із можливістю її постійного вдосконалення й розширення для включення нових багатопрофільних даних.

## РОЗДІЛ 2

### МАТЕРІАЛИ І МЕТОДИ

Матеріалом для виконання дисертаційної роботи слугували дані з генної експресії в плаценті й матково-плацентарному комплексі людини, у крові матері і плоду, які було отримано із баз ArrayExpress і GEO.

#### 2.1. Збір первинних даних

Збір первинних даних провели через звернення до бази ArrayExpress із пошуковими запитами, які включали комбінації термінів «placenta», «first trimester», «second trimester», «third trimester», «preeclampsia», «pre-eclampsia», «preeclamptic», відфільтрованих за організмом «Homo sapiens», типом експерименту «rna assay» та «array assay». Унаслідок пошуку ми обрали 43 датасети, які містили інформацію про експресію генів у плаценті, компонентах матково-плацентарного комплексу, плацентарних клітинах і в крові матері й дитини за умов фізіологічного й патологічного перебігу вагітності. Базу GEO використовували як додаткове джерело даних із генної експресії. Кілька датасетів були виключені з подальшого аналізу: E-GEOD-25906 у зв'язку з технічною помилкою під час завантаження даних, E-MTAB-3732 через те, що він є лише компіляцією датасетів, уже наявних на ArrayExpress, і дублює їх, а E-GEOD-15787, E-GEOD-22526 і E-MEXP-1050 були виключені через застарілі платформи мікромасивів і брак нуклеотидної послідовності для проб. Повний список включених та виключених наборів даних міститься в Додатку Е.

## **2.2. Створення спеціалізованої бази даних із генної експресії в плаценті та компонентах матково-плацентарного комплексу**

За зібраними даними генної експресії в плаценті та компонентах матково-плацентарного комплексу ми розробили спеціалізовану реляційну базу даних **IGEA** (Integrative Gene Expression Analysis), яка містить метадані для відповідних даних генної експресії. Метадані для всіх даних генної експресії були автоматично отримані з баз ArrayExpress та GEO. Тексти відповідних статей та листування з авторами були використані для отримання додаткових метаданих. Для стандартизації термінології в метаданих були використані Medical Subject Headings (MeSH) та Experimental Factor Ontology (EFO). MeSH використовується для кодифікації ключових слів у біологічних наукових статтях, а EFO — це онтологія, яку ArrayExpress використовує для текстового пошуку під час обробки пошукових запитів користувачів. Програмне забезпечення для створення спеціалізованої бази написано мовою Python з використанням фреймворку Django для розробки веб інтерфейсу та Postgres для підтримання бази даних. Початковий код та резервний файл бази даних **IGEA** доступні на GitHub за посиланням <https://github.com/Sashkow/placenta-preeclampsia>, а веб інтерфейс бази даних доступний за посиланням <https://molecularbiology.ink/igea/>. Етапи розробки бази даних **IGEA** наведено в Додатку Б.

## **2.3. Обробка первинних даних генної експресії, встановлення відповідності між пробкетами та генами і визначення показника експресії для кожного гена**

Первинні дані генної експресії з платформи Affymetrix представлені у вигляді .CEL файлів (cell intensity file), які містять результати обчислення

інтенсивності випромінювання порізнених пікселів. Для обробки цих даних ми використали пакет R `affy` [41]. Цей пакет містить функцію `rma` (robust multiarray average) [42], яка приймає на вхід `.CEL` файли та відповідний `CDF` файл (Content Definition File) анотації пробсетів, завантажений із сайту Brainarray, у якому встановлена взаємно однозначна відповідність між пробсетами й генами для конкретної моделі чипа Affymetrix. Виконання функції `rma` складається з декількох етапів. На початку встановлюється відповідність між пробами на мікрмасиві та генами. Після цього первинні значення інтенсивності випромінювання проб для кожного гена перетворюються у єдине значення експресії гена. Далі виконується логарифмування значень експресії, квантильна нормалізація та вилучається фонові технічна варіативність. У такий спосіб отримується *матриця* експресії генів із чипів Affymetrix.

Для обробки даних експресії з платформи Illumina ми використали пакет R `lumi` [43], у якому виконуються аналогічні кроки, як і для обробки даних із платформи Affymetrix. Єдиною суттєвою відмінністю є те, що для мікрмасивів Illumina у якості значення експресії гена ми згідно з [37, 38] (див. розділ літогляду 1.4) обирали пробсет із максимальним середнім значенням експресії генів серед усіх пробсетів, пов'язаних із цим геном і для всіх зразків у датасеті.

**2.3.1. Перевірка якості обробки даних.** Якість обробки даних і видалення ефекту серійності (batch effect) перевіряли за допомогою аналізу головних компонент (PCA, функція `prcomp` в R) [42] та *t*-розподіленого вкладення стохастичної близькості (англ. *t-distributed Stochastic Neighbor Embedding*, t-SNE: TSNE, пакет `Rtsne` в R) [44, 45]. Обидва методи є методами зменшення розмірності даних і їхньої візуалізації [45]. Внаслідок застосування цих методів первинні багатовимірні значення експресії генів відображаються у вигляді точок у двовимірному або тривимірному просторі.

## **2.4. Нормалізація й інтеграція даних експресії генів в однотипних датасетах**

Для об'єднання в одну групу однотипних (з однаковими метаданими) даних експресії з декількох датасетів необхідно усунути систематичні відмінності в даних, що виникають через небіологічні фактори, такі як варіації експериментальних умов або обробки зразків, а не через досліджувані біологічні фактори. Ця процедура зветься усуненням ефекту серійності (англ. Batch effect). Ми використали емпіричний метод Байеса — процедуру статистичного оцінювання, де апріорний розподіл значень експресії для кожного гена розраховується зважаючи на усереднений розподіл значень експресії генів у датасеті. Функція ComBat з пакету R limma [46] реалізує емпіричний метод Байеса. Для усунення ефекту серійності треба враховувати стать плоду у якості прихованої змінної (confounding variable) для того, щоби варіативність, пов'язана зі статтю, не була вилучена в процесі видалення ефекту серійності. Для зразків, у яких немає даних щодо статі плоду, ми визначили цей параметр за даними експресії генів, зчеплених з Y-хромосоною, і за допомогою пакету R massiR [47]. Більш детально процедуру описано у підрозділі 1.4.

## **2.5. Аналіз інтегрованих даних**

**2.5.1. Визначення диференційно експресованих генів.** Спрощено, можна сказати, що ген є диференційно експресованим, якщо є статистично достовірною різниця між, наприклад, середнім значенням його експресії для зразків першого триместру і для зразків другого триместру. Для визначення статистичної достовірності диференційної експресії ми застосували узагальнені лінійні моделі, реалізовані в R пакеті limma і визначили диференційно експресовані гени. Для кожного гена були підібрані параметри лінійної моделі та обчислені статистичні

характеристики за допомогою модерованої t-статистики. Скориговане р-значення (false discovery rate, FDR) визначено з урахуванням множинних порівнянь за методом Бенджаміні-Хошберга [48, 49].

**2.5.2. Пошук функціонально збагачених груп диференційно експресованих генів.** Щоб навести біологічну інтерпретацію знайдених диференційно експресованих генів, ми визначили функціонально збагачені групи генів. Спочатку ми побудували граф асоціації і взаємодії між білками, кодованими диференційно експресованими генами, використовуючи дані з бази String [50], а потім провели кластеризацію й візуалізацію графа, використавши пакети R STRINGdb і алгоритм fastgreedy з пакету R igraph. Вершинами графа є білки, що кодуються диференційно експресованими генами, а ребрами — безпосередні чи опосередковані білок-білкові взаємодії. Кластер — це група вершин графа, які між собою мають багато зв'язків і водночас мало зв'язків із рештою графа, що відображає їхню високу внутрішню згуртованість і розглядається як окрема функціональна або структурна одиниця графа. В отриманих кластерах (синоніми: спільнотах, модулях) ми визначили групи генів, які належать до певного біологічного процесу і зустрічаються в сукупності генів кластера частіше, ніж вони зустрічаються в геномі, тобто є збагаченими. Біологічний процес, до якого належать ці гени, має назву «функціонально збагачений біологічний процес (ФЗБП) (англ. functionally enriched biological process) [51]. ФЗБП ми визначили за даними з бази Gene Ontology і за допомогою гіпергеометричного тесту [52]. Він застосовується, коли потрібно перевірити, чи певна група об'єктів у вибірці відрізняється від очікуваного розподілу в популяції.

**2.5.2.1. Оптимізація побудови графа.** Кожна взаємодія в базі даних String оцінюється значенням достовірності (confidence), дійсним числом між 0 та 1, відповідно до сили взаємодії та надійності результатів досліджень. Якщо,

будуючи граф, обрати високий рівень достовірності, то дістанемо малі кластери з високою специфічністю та багато ізольованих вершин. Якщо ж обрати низький рівень достовірності, то дістанемо великі кластери з низькою специфічністю й навіть, можливо, без жодного збагаченого біологічного процесу. Відтак, ми ввели метрику *покриття графа біологічними процесами*, означивши її як частку генів у графі, які належать до одного або більше біологічних процесів, і обрали рівень достовірності рівним або вище 0,1, щоб максимізувати покриття графу біологічними процесами.

Для наочності для біологічних процесів у кожному кластері були побудовані теплові карти за допомогою пакета R *heatmap*, які показують попарний рівень близькості між біологічними процесами. За критерій близькості для пари біологічних процесів було взято кількість спільних диференційно експресованих генів між процесами у відношенні до кількості диференційно експресованих генів, які входять принаймні у один із процесів (кількість елементів у перетині множин поділити на кількість елементів у об'єднанні множин). Цей критерій відомий як індекс Джаккарда (*Jaccard index*) [53], і використовується для оцінки подібності двох множин, в тому числі множин функціонально збагачених генів [54]. Інтервал індексу Джаккарда від нуля до одиниці. Чим ближче схожість між збагаченими біологічними процесами, тим більше наближається індекс до одиниці.

**Результати даного розділу опубліковані в наступних роботах:**

Lykhenko, O., Frolova, A., & Obolenska, M. (2017) "Designing the database for microarray experiments metadata", *2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF)*. IEEE. Available at: doi:10.1109/YSF.2017.8126658.

## РОЗДІЛ 3

### РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

#### 3.1. Структура бази даних IGEA

Структура бази даних IGEA розроблена у такий спосіб, щоб зберігати в ній метадані про: біологічні зразки, дослідження та платформи мікромасивів; початкові назви і значення характеристик зразків (`old_name`, `old_value`) у тому вигляді, у якому вони були завантажені з ArrayExpress чи інших джерел; стандартизовані назви і значення характеристик (`name`, `value`), отримані відповідно до онтологій MeSH і EFO, й інформацію про те, яким стандартизованим назвам відповідають які стандартизовані значення. Наприклад, назві характеристики Biological Specimen відповідають значення Placenta, Chorion, Chorionic villi, Decidua, Maternal Blood, а назві Diagnosis — значення Healthy, Preeclampsia, Fetal Growth Retardation, HELLP syndrome.

Рис. 3.1 ілюструє загальну структуру бази даних. Прямокутникам відповідають таблиці бази. В таблицях Microarray, Experiment, Sample і SampleAttribute `id` означають номери відповідно мікромасивів, експериментів, зразків і характеристик зразків. Під терміном “data” в таблицях Microarray і Experiment містяться характеристики мікромасивів (тип платформи, ідентифікатор платформи) і експериментів (дата проведення, кількість зразків тощо). В таблиці SampleAttribute пари `old_name` і `old_value` і `name` і `value` містять текстову інформацію про початкові і стандартизовані назви і значення характеристик зразків, відповідно. Data та `additional_info` в таблицях Microarray, Experiment, StandardSampleAttributeName та StandardSampleAttributeValue містять списки пар типу «назва: значення».

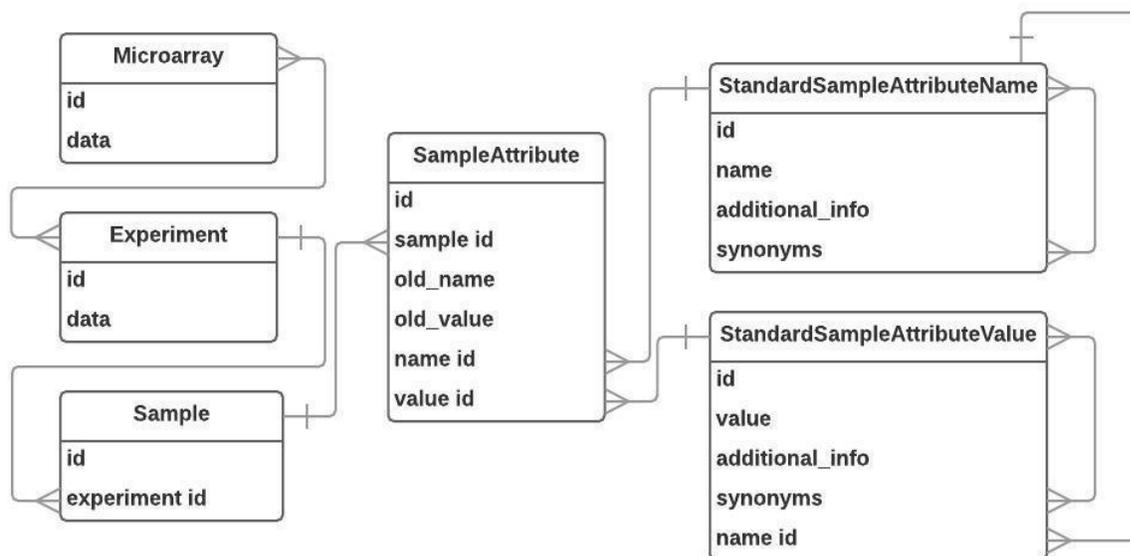


Рис. 3.1. Структура реляційної бази даних IGEA.

Примітка: id — ідентифікатор; old name і old value — початкові назви і значення характеристик зразків, завантажені з відкритих баз даних.

На схемі зв'язки між таблицями позначені стрілками, що ведуть від таблиці, яка посилається (наприклад, Sample, три розгалужені лінії на кінці стрілки), до таблиці, на яку посилаються (наприклад, Experiment, кінець стрілки перекреслений однією рисою). Такий тип взаємовідношення зветься один-до-багатьох, бо єдине унікальне значення id в колонці Experiment може зустрічатися кілька разів у (колонці) experiment id таблиці Sample, тобто в одному експерименті можуть бути залученими декілька зразків.

### 3.2. Класифікація зразків плаценти й матково-плацентарного комплексу за ключовими метаданими

На сьогодні в базу IGEA внесено дані з 48 датасетів (1466 зразків) матково-плацентарного комплексу та крові матері і плоду. Ми визначили три ключові

характеристики, за якими класифікували зразки: діагноз, строк гестації й тип біологічного матеріалу (ворсини хоріону, плацента, децидуа, амніон, клітини плаценти, кров матері й кров плоду) (рис. 3.2).



Рис. 3.2. Класифікація зразків матково-плацентарного комплексу за ключовими метаданими

**Строк гестації** часто не вказується в метаданих зразка, і це дуже прикро. Проте його можна приблизно визначити, наприклад, якщо в початкових даних зазначено, що пологи відбулися вчасно (37–41 тиждень вагітності) або передчасно (20–37 тиждень). Також автори іноді наводять значення віку плоду, чого цілком достатньо, щоби віднести зразки до певної категорії за строком гестації. Строк гестації або вік плоду наводять у днях, тижнях і місяцях.

За строком гестації зразки, занесені до бази даних IGEA, розподіляються на 5 груп: дві групи зразків належать до першого (1–12 тиждень) і другого (13–24 тижні) триместру, дві групи — до ранніх (25–30 тижнів) і пізніх (30–37 тижнів) передчасних пологів і одна група — до вчасних (37–41 тиждень) пологів.

**Діагноз** вказується майже в усіх первинних даних, але часто значення цього атрибута є недостатнім. Наприклад, якщо зразок належить до контрольної групи,

це може бути плацента, отримана внаслідок природних пологів або внаслідок кесаревого розтину; пологи можуть бути вчасними або передчасними; індукованими або спонтанними. Усі ці характеристики мають бути враховані під час об'єднання даних у групи випадок — контроль. Так само, якщо в діагнозі вказано термін «пreekлампсія», вона може бути з ранніми або пізніми симптомами, різної тяжкості, додатковими ускладненнями тощо.

**Тип тканини.** У третині випадків вказується лише орган — плацента. Втім, через тканино- і клітинспецифічні маркери експресії генів можна визначити до якого компонента матково-плацентарного комплексу належить зразок.

Іноді метадані вдається доповнити завдяки прочитанню текстів статей, листування з авторами, а також у деяких випадках, використовуючи дані генної експресії, як, наприклад, із визначенням статі плоду в [55] (рис. 3.3).

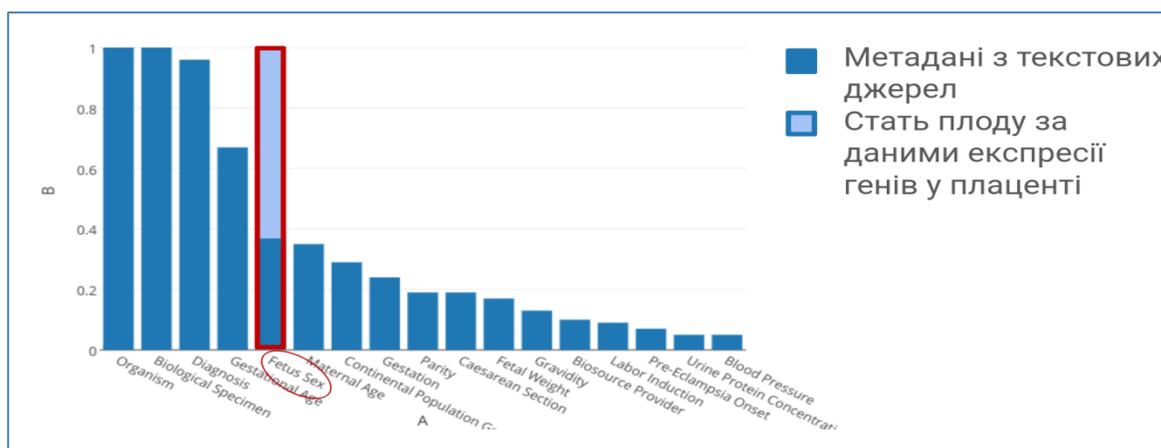


Рис. 3.3. Розподіл зразків плаценти за метаданими

### Результати підрозділів 3.1 і 3.2 опубліковані в наступних роботах:

Lykhenko, O., Frolova, A., & Obolenskaya, M.(2017). Creation of gene expression database on preeclampsia-affected human placenta. *Biopolym. Cell*, 33(6), 442–452. doi: 10.7124/bc.000967

Lykhenko, O., Frolova, A., & Obolenska, M. (2017) "Designing the database for microarray experiments metadata", *2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF)*. IEEE. Available at: doi:10.1109/YSF.2017.8126658.

Frolova A, Lykhenko O, Bondarenko V., and Obolenskaya M. (2017(2019)). Step by step to integration of gene expression profiles for preeclampsia-affected human placenta. China — Ukraine International Symposium on Innovation and Technology, (2017). 30.10 — 03.11. 2017. Shandong Academy of Sciences. Abstract book. P. 1

Lykhenko, O., Frolova, A., & Obolenskaya, M. (2021). Consecutive integration of available microarray data for analysis of differential gene expression in human placenta. *Biotechnologia Acta* 14(1), 38-45. Available at: doi: <https://doi.org/10.15407/biotech14.01.38>

### **3.3. Загальна характеристика генної експресії між першим і другим та другим і третім триместрами нормальної вагітності**

Послугуючись стандартизованими метаданими IGEA, з відкритих з баз GEO і ArrayExpress ми завантажили дані з 8 досліджень генної експресії в тканині плаценті з трьох триместрів нормального перебігу вагітності. З розробленої нами бази даних IGEA ми завантажили стандартизовані метадані з восьми досліджень зразків плаценти, строк гестації – I, II і III триместр, діагноз – нормальна вагітність, а з баз GEO і ArrayExpress ми завантажили відповідні дані генної експресії з тих самих восьми досліджень (табл. 3.1).

Таблиця 3.1

## Структура даних генної експресії в плаценті людини

Датасети (база GEO)	Кількість зразків						Кількість генів
	Загальна	За триместрами			За статтю плоду		
		I	II	III	Жіноча	Чоловіча	
GSE122214	4	4			2	2	20261
GSE22490	6	5	1		4	2	20261
GSE35574	26			26	10	16	24547
GSE37901	4		4		1	3	20342
GSE6573	1			1	0	1	20326
GSE73374	12			22	8	4	29509
GSE73685	12			12	12	0	20624
GSE9984	12	4	4	4	9	3	20342
<b>Загальна кількість</b>	<b>77</b>	<b>13</b>	<b>9</b>	<b>55</b>	<b>46</b>	<b>31</b>	

## 3.4 Оцінювання якості видалення ефекту серійності

Під час інтеграції даних з декількох досліджень ми усували ефект серійності, тобто технічної, небіологічної варіативності даних. Під час усунення ефекту серійності важливо, щоб збереглися відмінності, які зумовлені біологічною природою явища і процесу. Для перевірки якості видалення ефекту серійності ми використали аналіз головних компонент (різновид методу машинного навчання без учителя), який надає первинні багатовимірні дані у двовимірному просторі. На рис. 3.4 (а-д) наведений приклад аналізу головних компонент для перевірки видалення ефекту серійності під час інтеграції даних генної експресії з чотирьох датасетів у першому та другому триместрах здорової вагітності. Рис. 3.4а зображує розподіл даних перед видаленням ефекту серійності. Видно, що за

генною експресією найбільш відрізняються між собою зразки з датасетів GSE122214 і GSE9984 і зразки кожного з них від інших двох (принципові компоненти  $\text{dim1}$  та  $\text{dim2}$ ), хоча всі дані експресії належать до першого і другого триместрів. Найбільшу варіативність (принципові компоненти  $\text{dim1}$  та  $\text{dim2}$ ) вносять різні датасети. Ця технічна варіативність зникає після видалення ефекту серійності (рис. 3.4б). На противагу технічній варіативності, різниця між зразками першого та другого триместру не була прихована перед видаленням ефекту серійності (рис. 3.4в), але стала основним джерелом варіативності після його видалення (рис. 3.4г). Дані генної експресії можна розділити й за статтю плоду, що представлено в компонентах  $\text{dim3}$  та  $\text{dim4}$  (рис. 3.4д).

У незначній кількості випадків ми усували ефект серійності й у межах датасету одного дослідження, коли метод головних компонент показував неочікуване групування зразків, яке ми не могли пояснити біологічно, але нам вдалося знайти технічну причину варіативності даних експресії. В одному випадку грав роль день експерименту, зазначений в метаданих, тобто дані експресії були систематично подібними буквально в межах серії зразків. В іншому випадку чіп мікромасива складався з двох «половинок» (Affymetrix HG-U133A і Affymetrix HG-U133B) і в цьому випадку ефект серійності стосувався не груп зразків, а груп генів, чії проби розташовані на тій чи іншій «половинці» чіпа.

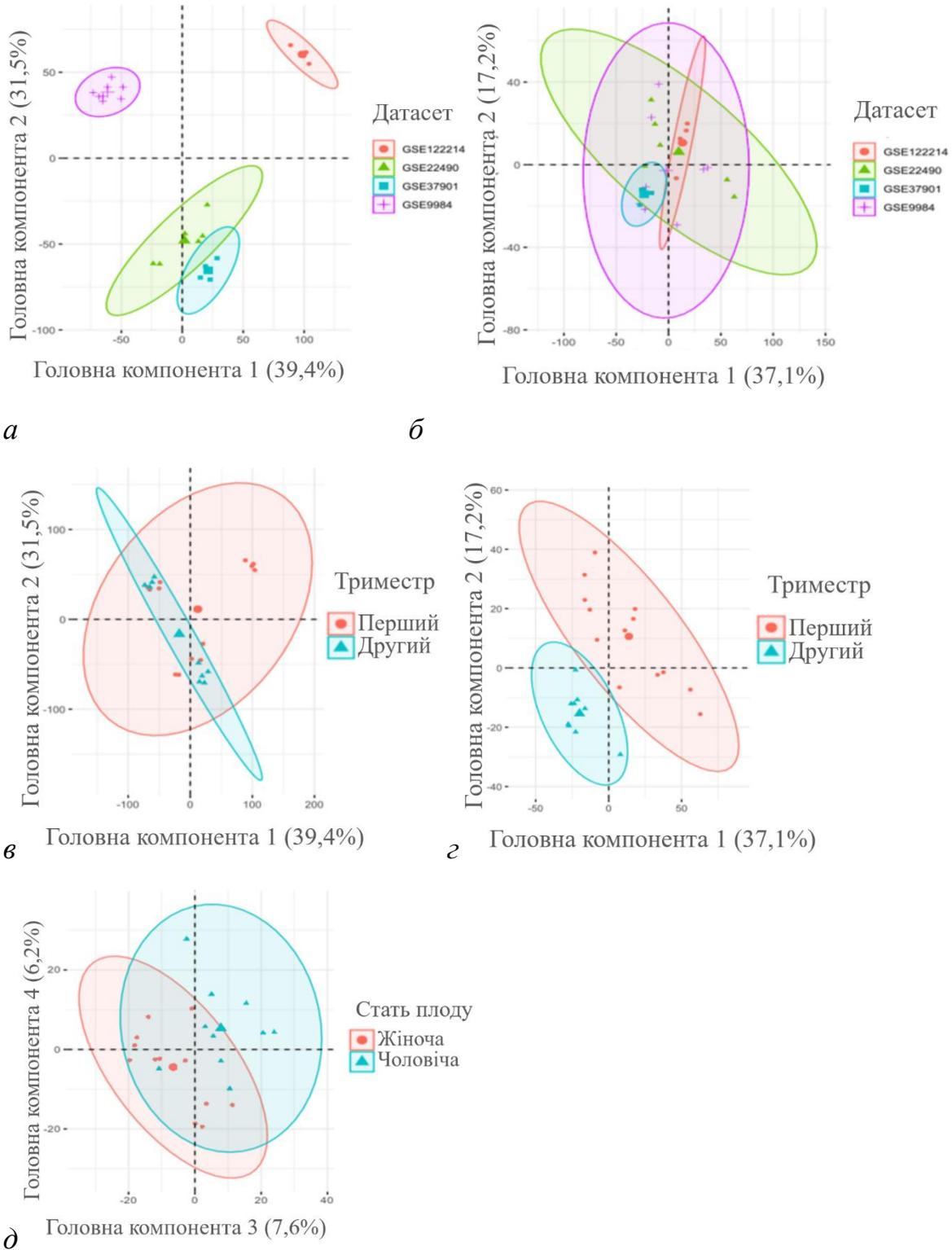


Рис. 3.4. Результати аналізу головних компонент до (3а і 3в) й після (3б, 3г і 3д) видалення ефекту серійності

### **3.5. Визначення і характеристика диференційно експресованих генів між триместрами**

Визначення генної експресії у восьми дослідженнях базувалося на реакції гібридизації відповідних генам людини коротких кДНК олігонуклеотидів (25 і 50 нуклеотидів відповідно для чипів Affymetrix та Illumina) на чипі мікромасива з кДНК, отриманою в результаті обробки мРНК зі зразків плаценти. Рівень гібридизації свідчить про те, скільки тієї чи іншої унікальної мРНК міститься в препараті сумарної мРНК, і віддзеркалює експресію білок-кодуєчих генів (надалі по тексту – транскриптом).

Ми провели обробку первинних даних для кожного триместру і попарно об'єднали дані експресії генів із першого і другого триместру (порівняння 1\_2) та з другого і третього триместру (порівняння 2\_3). У кожній парі визначили: диференційно експресовані гени (ДЕГи) і їх належність до певного кластера за методом fastgreedy графі, побудованого за даними бази String, як вказано у підрозділі 2.5. У цій базі інтегровані відомі й передбачені міжбілкові фізичні взаємодії та функціональні асоціації. Для цього зібрані й оцінені у балах дані з різних джерел як то: автоматичний аналіз наукової літератури; база експериментальних даних щодо міжбілкової взаємодії та анотованих міжбілкових комплексів/шляхів; комп'ютерні передбачення щодо взаємодії на підставі ко-експресії генів і консервативності геномного контексту, і відомості щодо міжбілкових взаємодій у різних організмів [50]. У найбільших за чисельністю генів кластерах визначили категорії біологічних процесів за допомогою програми Stringdb. Кластери названі відповідно до найбільш представлених у них біологічних процесів.

Серед ДЕГів у порівнянні першого триместру з другим (порівняння 1\_2) за допомогою алгоритму fastgreedy виявили 8 кластерів у порівнянні першого з

другим триместром і 14 кластерів у порівнянні другого з третім триместром. Серед них із величиною покриття збагачення (enrichment coverage) у межах 0.87 – 1.0 виявили 4 кластери у порівнянні 1\_2 і 7 кластерів з величиною покриття 0.14 – 0.95 у порівнянні 2\_3 (табл. 3.2 і 3.3, рис. 3.5 і 3.6). Величина покриття збагачення (enrichment coverage) вимірюється у межах 0 – 1 і відображає частку генів кластера, які належать до принаймні однієї функціонально збагаченої групи генів за даними програми Stringdb.

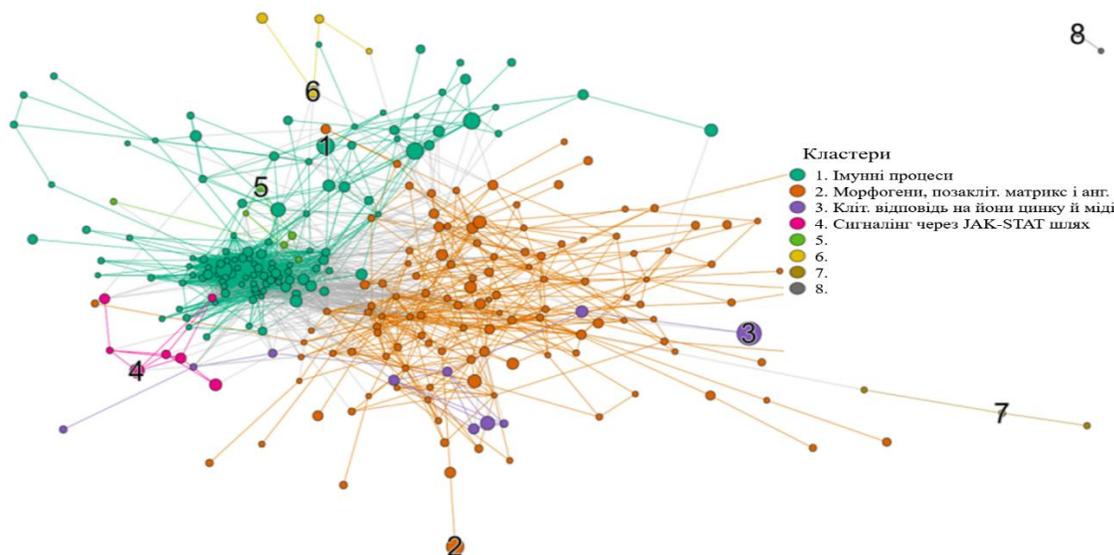


Рис. 3.5. Граф білок-білкових взаємодій та асоціацій згідно переліку диференційно експресованих генів між другим і першим триместрами вагітності

Примітка: кожний кластер генів з його вершинами і ребрами позначений своїм кольором. Вершини і ребра, які не увійшли до кластерів, позначені сірим кольором. Розмір вузлів (вершин) віддзеркалює інтенсивність експресії ДЕГів генів між I і II триместрами. Кластери 5-8 не мають назви, бо для них не було визначено жодного збагаченого біологічного процесу.

Таблиця 3.2

**Загальна характеристика кластерів ДЕГів між першим і другим триместрами вагітності (порівняння 1\_2)**

Кластери	Кількість генів			logFC мін.	logFC макс.	Покриття
	Загальна	з напрямком змін				
		угору	униз			
1. Імунні процеси	114	81	33	-3,23	2,97	0,87
2. Морфогени, позаклітинний матрикс і ангиогенез	122	62	60	-3,2	2,15	0,88
3. Клітинна відповідь на йони цинку й міді	10	3	7	-4,43	1,55	1,0
4. Сигналінг через JAK-STAT шлях	7	6	1	-2,68	2,25	1,0
<b>Усього</b>	<b>253</b>	<b>152</b>	<b>101</b>			

Примітка: напрямок змін «угору» означає підвищення експресії диференційно експресованих генів із часом гестації, а напрямок змін «униз» означає зниження експресії з часом гестації. logFC мін. — мінімальне значення логарифму кратності змін; logFC макс. — максимальне значення логарифму кратності змін.

Ми визначили 268 генів, які диференційно експресовані у порівнянні 1\_2, і 489 генів у порівнянні 2\_3, що майже вдвічі перевищує кількість диференційно експресованих генів у попередній період. Ця різниця між двома порівняннями може бути частково пов'язана з різним часовим проміжком у двох порівняннях: у порівнянні 1\_2 він складає від 6 до 11 тижнів (I триместр, 7 – 11 тижнів і II

триместр, 13 – 18 тижнів вагітності). У порівнянні 2\_3 він складає від 24 до 25 тижнів (II триместр, 13 – 18 тижнів і III триместр, 38 – 42 тижня).

Кластери, пов'язані з імунними процесами і з морфогенезом і ангиогенезом є найбільш чисельними впродовж переходу від першого до другого триместру.

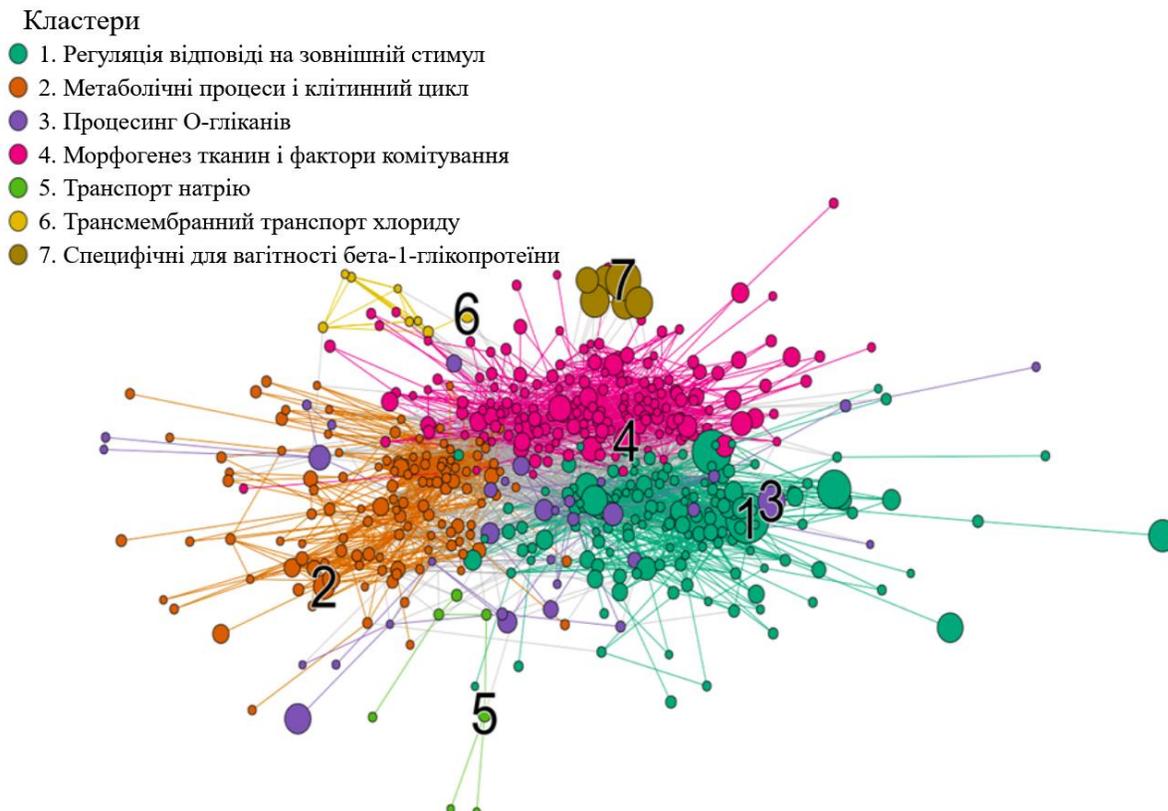


Рис. 3.6. Граф білок-білкових взаємодій та асоціацій відповідно до диференційно експресованих генів між другим і третім триместрами (порівняння 2\_3). Примітка: дивись позначки до рис. 3.5.

Період 2\_3 у порівнянні з 1\_2 інтервалом характеризується більшою кількістю кластерів диференційно експресованих генів і кількістю ДЕГів.

Таблиця 3.3

**Загальна характеристика кластерів ДЕГів між другим і третім триместрами вагітності (порівняння 2\_3)**

Кластери	Кількість генів			logFC мін.	FC макс.	Покриття збагачення
	Загальна	з напрямком змін				
		угору	униз			
1. Регуляція відповіді на зовнішній стимул	145	56	89	-3,94	6,14	0,95
2. Метаболічні процеси і клітинний цикл	119	43	76	-1,91	2,97	0,93
3. Процесинг О-гліканів	36	24	12	-2,13	6,65	0,14
4. Морфогенез тканин і фактори комітування	168	84	84	-2,68	3,58	0,91
5. Транспорт натрію	7	3	4	-1,83	1,48	0,57
6. Трансмембранний транспорт хлориду	8	5	3	-2,01	1,44	1
7. Специфічні для вагітності бета-1-глікопротеїни	6	6	0	3,16	4,97	1
Усього	<b>489</b>	<b>221</b>	<b>268</b>			

Примітка: дивись позначки до табл. 3.2.

### **3.6. Кластери диференційно експресованих генів і функціонально збагачені біологічні процеси між першим і другим триместром**

В цьому і подальших розділах ми наводимо диференційно експресовані гени зі значенням логарифму кратності зміни експресії за основою 2 (logFC) у круглих дужках жирним шрифтом, наприклад, ***CCL2 (1,64)***. Повний перелік диференційно експресованих генів і функціонально збагачених біологічних процесів між

першим і другим триместром вагітності надані у Додатку В. Кожний із кластерів проілюстровано теплокартою функціонально збагачених біологічних процесів (ФЗБП) за близькістю генів в попарних процесах і блок-діаграмами найбільших за чисельністю ФЗБП і експресії генів в одному з процесів у другому триместрі у порівнянні з першим в якості прикладу.

**3.6.1. Кластер 1. Імунні процеси.** Вагітність є викликом для імунної системи матері, плаценти і плоду, оскільки плод, який розвивається всередині генетично відмінного організму матері, не має бути відторгнутим і одночасно має бути захищеним від дії патогенних чинників, які можуть потрапити з крові та матки матері [56].

Протягом гестації такі події, як імплантація та плацентація (розвиток плаценти), ріст плоду та пологи, потребують унікального імунного середовища впродовж кожного періоду вагітності. Перелік різномірних збагачених процесів наведений на теплокарті (рис. 3.7). Цей перелік засвідчує активацію прозапальних процесів, активне переміщення клітин, активацію синтезу цитокінів, активацію регуляції імунних клітин (лейкоцитів і дендритних клітин), реакцію на стрес та інші процеси. Процеси з найбільшим числом генів наведені на рис. 3.8а.

Імунні процеси проілюстровано експресією хемокінів, серед яких експресія прозапальних хемокінів підвищується під час переходу від першого до другого триместру, а експресія гомеостатичних хемокінів знижується (рис. 3.8б). Прозапальні хемокіни залучають імунні клітини до місць запалення та індукують вивільнення супероксидних іонів, що беруть участь у запаленні, проліферації, метаболізмі та ангиогенезі [57]. Окрім залучення імунних клітин, хемокіни підтримують ремоделювання тканини та сприяють ангиогенезу. Порушення у синтезі хемокінів можуть мати шкідливі наслідки для організму. Наприклад, зміни

в експресії *CCL2* (1,64) в трофобласті й децидуа асоціюються з ускладненнями вагітності прееклампсією (Lin et al., 2023).

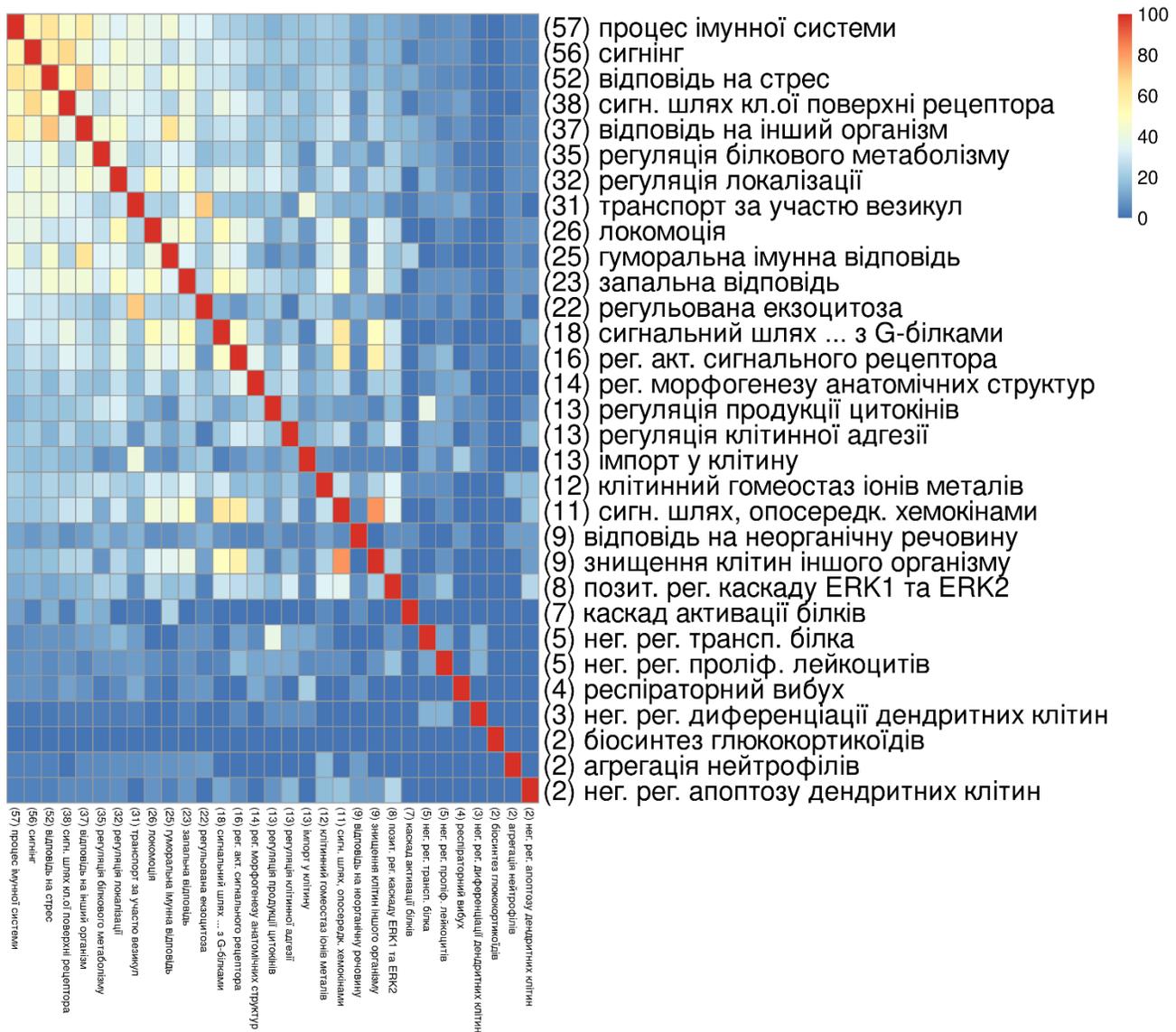


Рис. 3.7. Теплокарта функціонально збагачених біологічних процесів у кластері “Імунні процеси”

Примітка: Цифри в дужках відповідають кількості генів у ФЗБП. Колір на термошкالی відповідає попарній близькості процесів за кількістю ідентичних генів (у відсотках). У теплокарті ФЗБП за жовтим кольором, що переходить у коричневатий, виокремлюються ФЗБП, які поділяють із сусідами близько 60%



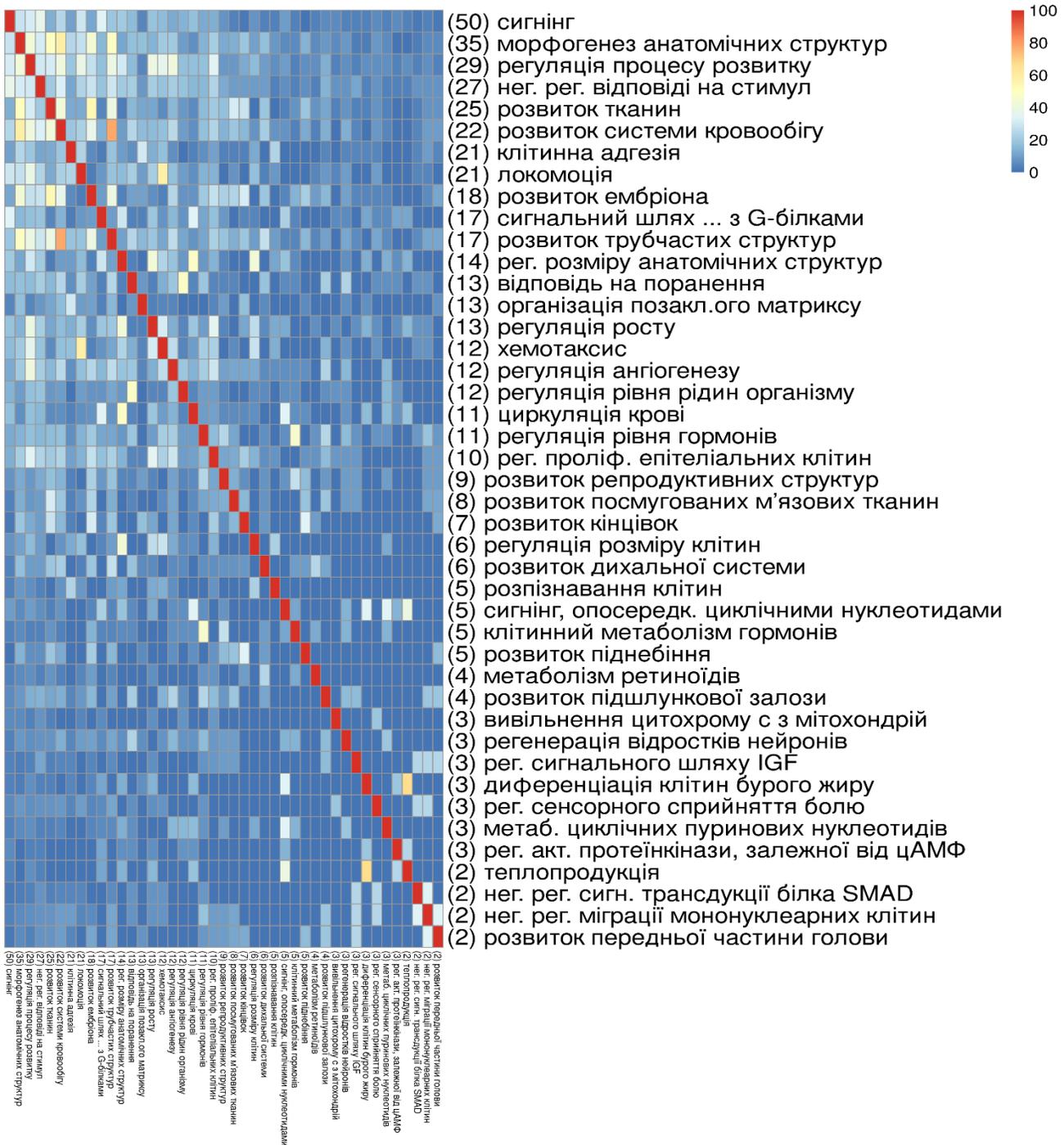


Рис. 3.9. Теплокарта функціонально збагачених біологічних процесів у кластері “Морфогени, позаклітинний матрикс і ангиогенез”

Див. примітки у підпису до рис. 3.7.

Участь диференційно експресованих генів у морфогенезі проілюстровано на прикладі сигналіngu через компоненти висококонсервативних Wnt-, Notch-, BMP-, Notch, і Hippo каскадів (Рис. 3.10), зайнятих у морфогенезі [58]. **Glypican-3 (GPC3)** — широко розповсюджений консервативний протеоглікан клітинної поверхні. Він складається з двох ланцюгів гепарансульфату, зв'язаних з коровим білком, який закорений у клітинній мембрані. Він може взаємодіяти з численними лігандами/морфогенами і тому відіграє значну роль у регуляції каскадів морфогенів, у рості та диференціюванні клітин, особливо в ембріогенезі.

Декілька генів є гомеотичними (Homeotic genes), тобто вони є транскрипційними факторами, які мають у своїй білковій послідовності гомеодомен, через який зв'язуються з регуляторними ділянками генів-мішеней і визначають інтенсивність їхньої експресії. Гомеотичні гени активно задіяні у розвиткові судинної системи і особливо у функціонуванні ендотеліальних клітин [59], але не тільки. Наприклад, гомеотичний ген **MEOX2**, мезодермальний фактор транскрипції, бере участь у сомітогенезі, розвитку склеротому і міогенезі кінцівок. Морфогени BMP — представники bone morphogenic proteins, найбільшої підродини (20 ізоформ) надродини трансформуючого фактору росту TGF- $\beta$ . Незважаючи на те, що свою назву гени отримали через їх роль у формуванні кісток і хрящів, представники цієї родини беруть участь у спеціалізації клітин, диференціюванні, проліферації і апоптозі, у визначенні осей тіла на ранньому етапі розвитку при формуванні організму і загалом у морфогенезі [60]. **BMP5** в N-глікозильованій формі сприяє проліферації, міграції, інвазії клітин трофобласта та ангиогенезу через BMP5-SMAD1/5 сигнальний шлях [61].

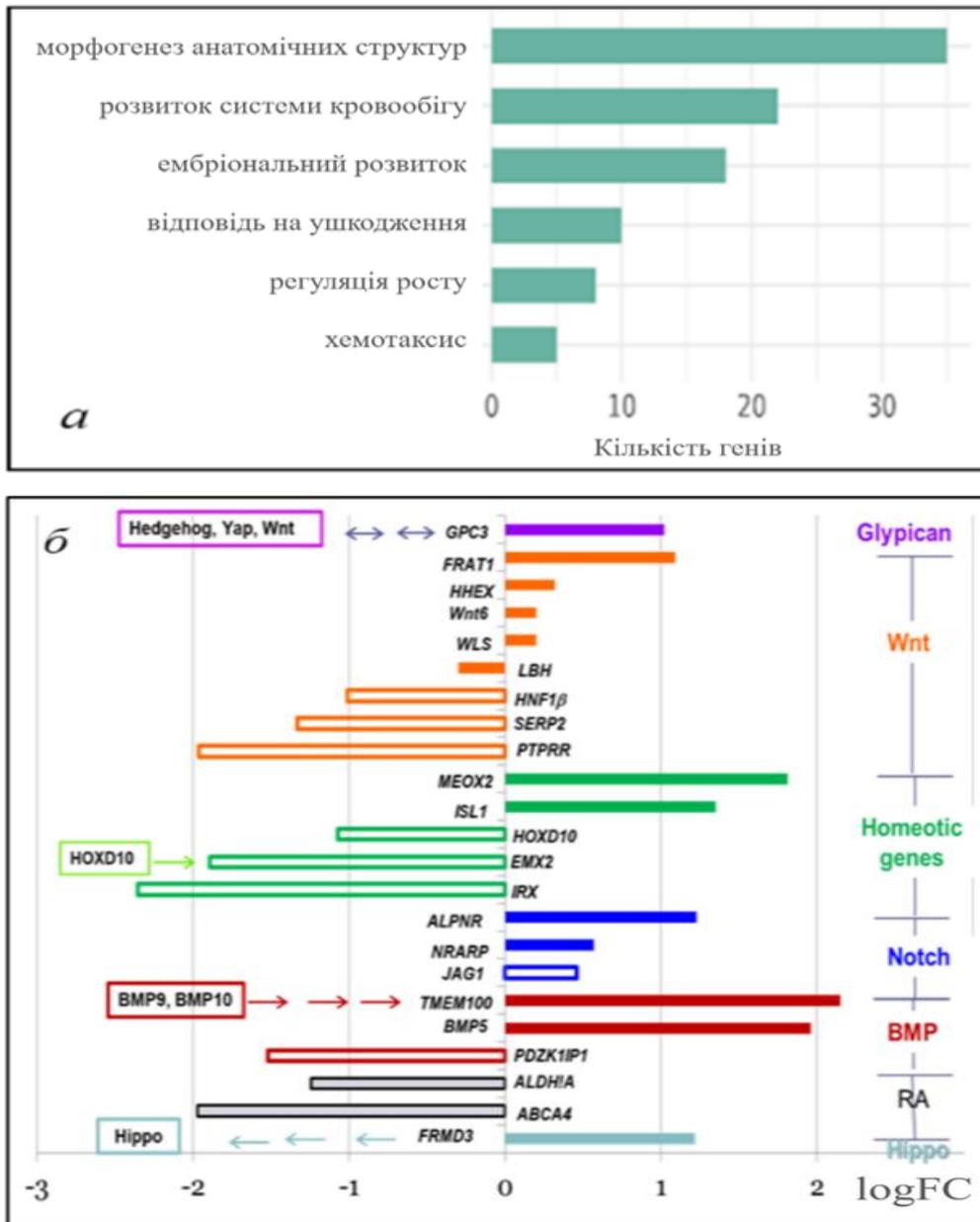


Рис. 3.10. Кількість генів для кожного з топ-6 функціонально збагачених біологічних процесів (а) і зміна експресії чинників морфогенезу у другому триместрі у порівнянні з першим у кластері “Морфогени, позаклітинний матрикс і ангиогенез” (б)

Примітка: гени-активатори на рис. 3.10б позначені зафарбованими блоками, а гени-інгібітори — незафарбованими блоками. Колір блока і контуру відповідає кольору назви каскаду, позначеного справа від діаграми. Ген EMX2 є мішенню

для транскрипційного фактору HOXD10; TMEM100 є мішенню у шляху сигналінгу BMP9 і BMP10; FRMD3 є вищим по течії (upstream) регулятором Hippo каскаду.

Ген *TMEM100* із підродини BMP важливий для диференціації артеріального ендотелію та судинного морфогенезу та діє у відповідному сигнальному шляху нижче за BMP9/BMP10-ALK1 [62]. А ген *PDZK1IP1* із тієї підродини перешкоджає утворенню комплексу R-Smad/Smad4, індукованому TGF- $\beta$  та BMP, і таким чином пригнічує цей шлях [63], але його експресія знижується. Як видно на рис. 3.10 у більшості випадків експресія генів-активаторів підвищується, а генів-інгібіторів синхронно знижується між першим і другим триместром, таким чином забезпечуючи максимальну ефективність шляху.

Ретиноева кислота (RA) критично необхідна на найранішніх стадіях ембріогенезу під час формуванн осей тіла, нейро- і кардіогенезу, розвитку підшлункової залози, легень і ока [64]. Надлишок ретиноевої кислоти виявляє тератогенний ефект. Припускаємо, що знижена експресія генів, причетних до синтезу ретиноевої кислоти, запобігає цьому ефекту і вказує на втрату їх необхідності.

Потрібно відмітити, що сукупність різних генів з різними функціями разом зумовлюють формування тканин і органів у ембріогенезі. Серед ФЗБП на рис. 3.9 привертає увагу декілька процесів, які позначені як відповідальні за розвиток органів і структур тіла (легенів, підшлункової залози, кінцівок, розвиток ембріону). Такими їх автоматично визначає система Gene Ontology, за якою названі ФЗБП. Наприклад, два набори генів, *FOXF2, MEOX2, TCF21, COL2A1, OSR2* і *COMP, HOXD10, MEOX2, SFRP2, COL2A1, GPC3, OSR2*. зумовлюють у ембріогенезі розвиток піднебіння і кінцівок, відповідно. Три гени є спільними між двома наборами, а інші ймовірно надають специфічності кожному з процесів

формування структур. Всі гени є білок-кодуєчими. Спільні гени: *MEOX2*, Mesenchyme Homeobox 2, мезенхімальний транскрипційний фактор, який інтенсивно експресується у плаценті і у серці плода; *OSR2*, транскрипційний фактор 2 (Odd-skipped related transcription factor 2) і *COL2A1*, який відповідає за синтез колагену II типу. Серед інших генів: *FOXF2*, транскрипційний фактор, активно експресується в плаценті, регулює проліферацію і стійкість до апоптозу через  $\beta$ -катенін та шлях Wnt [65, 66]; *TCF21*, повсюдний транскрипційний фактор, який відіграє важливу роль у специфікації долі та диференціюванні клітин різного типу. *TCF21* особливо інтенсивно експресується у плаценті і легенях; *COMP* пригнічує апоптоз, блокуючи активацію каспази-3 та індукуючи родину білків виживання IAP (BIRC3, BIRC2, BIRC5 та XIAP); *HOXD10*, гомеотичний транскрипційний фактор, є учасником регуляторної системи розвитку, що позиціонує клітини на передньо-задній осі; *SFRP2* (Secreted frizzled-related protein 2) -розчинний медіатор Wnt сигналіngu, показник пухлинного росту, і *GPC3* (гліпікан-3) бере участь у декількох сигнальних шляхах і виявляє шлях-специфічну активність у процесах росту та розвитку клітин/тканин. Таким чином, численні гени у цих двох вибірках кодуєють транскрипційні фактори, що зумовлює широкий спектр біологічної активності двох означених груп. За своєю здатністю зумовлювати долю і спеціалізацію клітин їх можна віднести до факторів комітування. Їх роль у розвитку певних структур і органів ембріона пов'язана зі специфічним набором діючих факторів, конкретним позиціонуванням процесу, часовим періодом його реалізації в процесі розвитку і специфічною інтенсивністю експресії генів. Молекулярний аспект їх участі у розвитку плаценти вивчений недостатньо.

**3.6.3. Кластер 3. Клітинна відповідь на йони цинку й міді.** ФЗБП кластера представлені на теплокарті (рис. 3.11). Кластер вирізняється зростаючою

активністю генів, які беруть участь у метаболізмі жирів (усі ДЕГи кластера зображено на рис. 3.12) (Ендотеліальна ліпаза, *LIPG*, 1,39; ліпопротеїнова ліпаза, *LPL*, 1,27 ікіназа піруватдегідрогенази 4, *PDK4*, 1,55). Перші дві ліпази розщеплюють відповідно фосфоліпіди і триглицериди, забезпечуючи тканину і плод вільними жирними кислотами і моноглицеридами як джерелом енергії [67, 68]. Білок **PDK4** у плаценті є домінуючим регулятором метаболізму глюкози та жирних кислот. Ця кіназа знижує активність піруватдегідрогеназного комплексу у мітохондріях шляхом його фосфорилування і тим самим сприяє анаеробному гліколізу, тобто окисненню пірувата у цитоплазмі з утворенням лактату і посилює синтез жирних кислот.

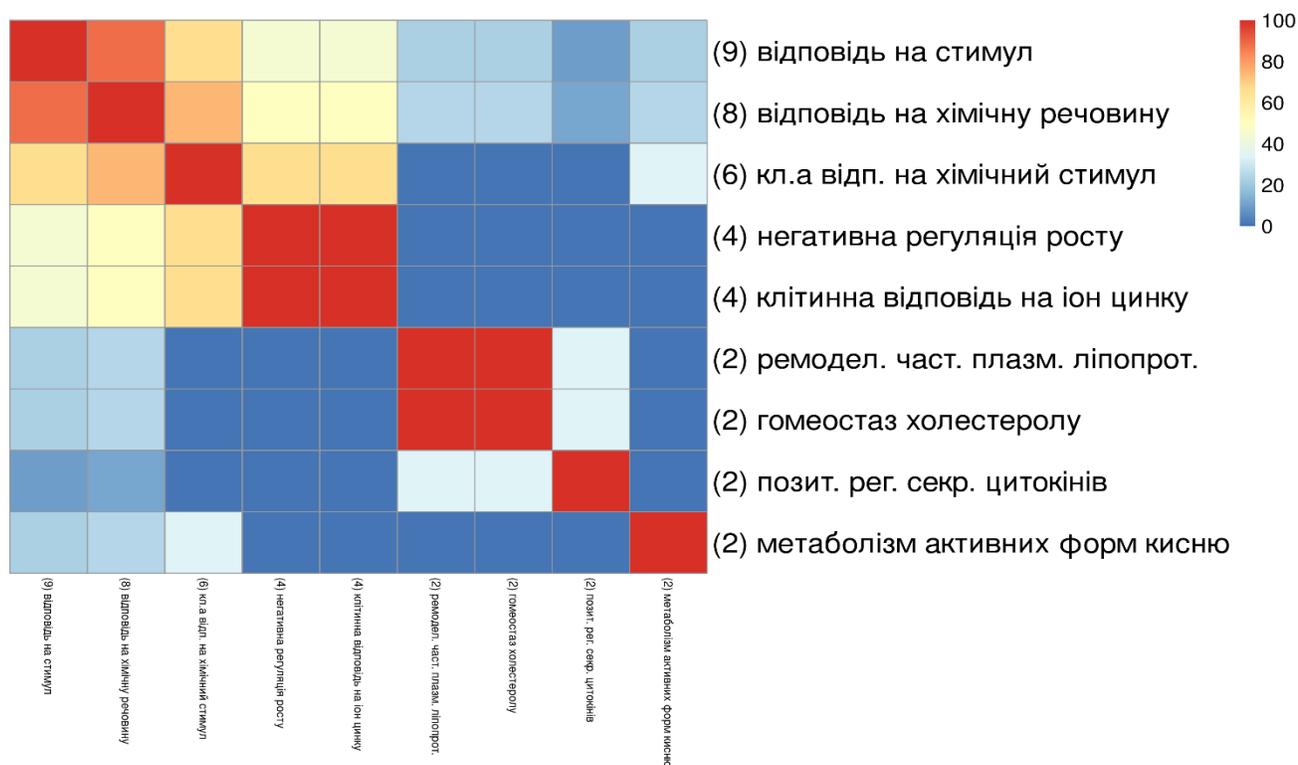


Рис. 3.11. Теплокарта збагачених біологічних процесів у кластері “Клітинна відповідь на йони цинку й міді”

Примітка: див. рис. 3.7.

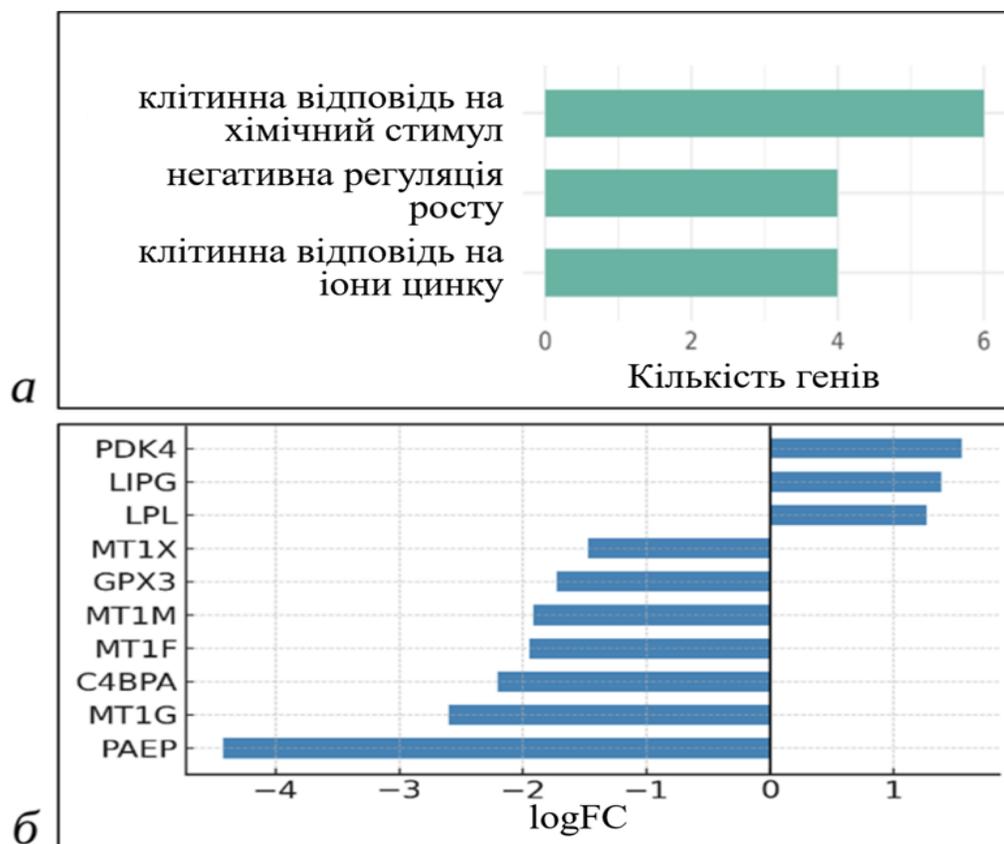


Рис. 3.12. Кількість ДЕГів для кожного з топ-3 функціонально збагачених біологічних процесів (а) і зміна експресії диференційно експресованих генів у другому триместрі у порівнянні з першим у кластері “Клітинна відповідь на йони цинку й міді” (б)

Знижується активність металотіонеїнів (*MT1*), які хелатують іони важких металів і зв'язують вільні радикали [69], і глутатіонпероксидази (*GPX3*), яка відновлює органічні пероксиди і перекис водню. Обидва зниження позбавляють захисту від оксидативного стресу у тканині. Іще знижується експресія прозапального гена *PAEP* (-4,43), який підтримує толерантність на етапі плацентації і інгібує активність імунних клітин і забезпечує прикріплення бластоцисти до стінки матки [70, 71], і гена-інгібітора у системі комплемента

*C4BPA* (-2,2), тим самим підсилюючи протизапальну активність системи комплементу.

**3.6.4. Кластер 4. Сигналінг через JAK-STAT шлях.** Гени цього кластера залучені до сигнального шляху, опосередкованого інтерлейкіном-35 (*EBI3*, 1,89; *IL12RB2*, 1,25; *CSF2RB*, 1,44) і до каскаду JAK-STAT (*GHI*, growth hormone 1,94; *CSH1*, chorionic somatomammotropin hormone 1 or placental lactogen, 1,64); *CSHL1*, chorionic somatomammotropin hormone like 1, 2,25) і *LEP* (-2,68) (рис. 3.13, 3.14).

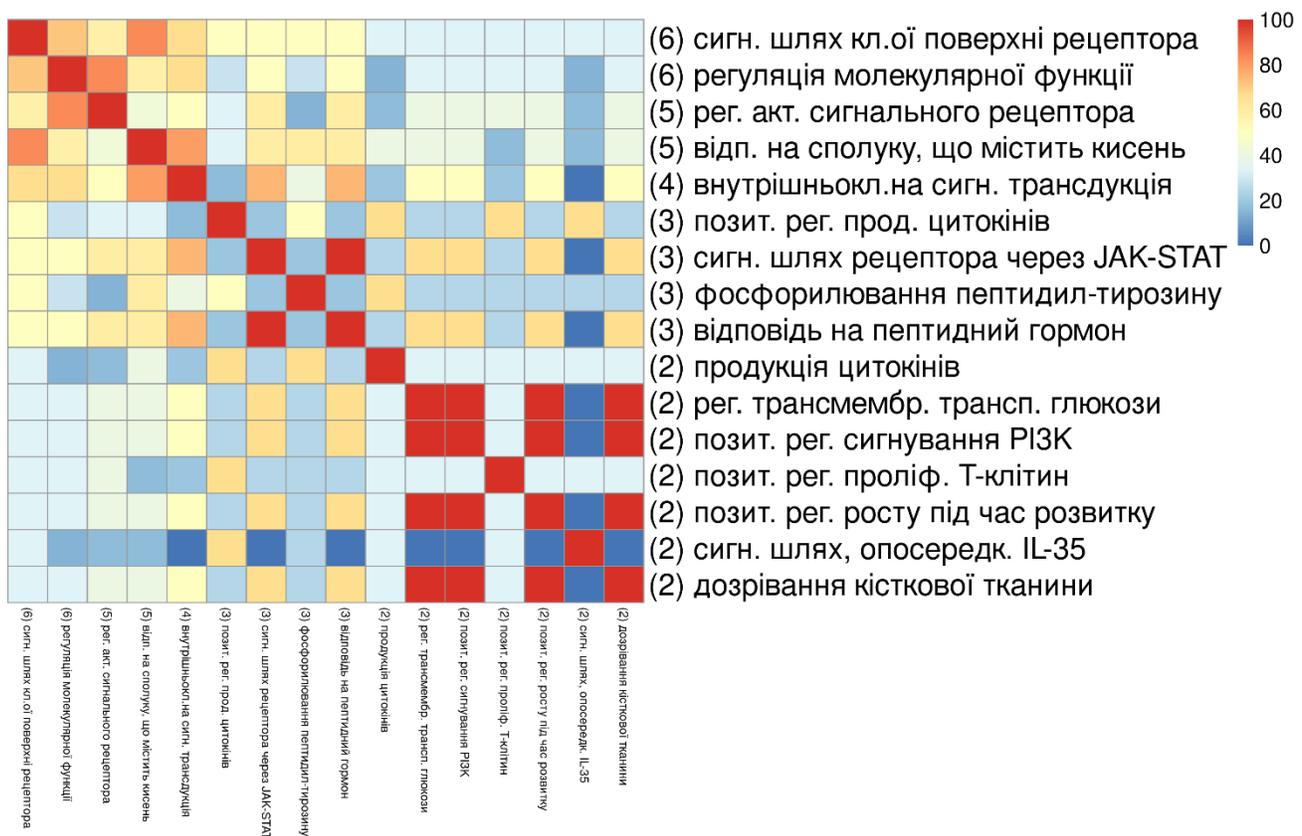


Рис. 3.13. Теплокарта збагачених біологічних процесів у кластері "Клітинна відповідь на йони цинку й міді"

Перша трійка генів залучена до активності цитокінів родини ІІ6/ІІ2 і колоній-стимулюючого фактору. Друга трійка генів кодує секреторні білки, за

своєю дією подібні до дії плацентарного гормону росту [72, 73]. В організмі матері **CSHL1** викликає помірковану інсулінорезистентність, що покращує надходження глюкози до плоду і активує синтез інсуліноподібного фактору росту [73].

Дія ростових гормонів перегукується з дією іншого білкового гормону, лептину (**LEP, -2,38**). Плацента займає друге місце після жирової тканини за інтенсивністю синтезу і постачання лептину в материнський кровообіг. Основна функція лептину — регуляція запасів жиру й сигналізація про відчуття голоду/ситості. Припускаємо, що зниження експресії гена лептину від I до II триместру може пояснювати, зокрема, підвищення апетиту на початку вагітності.

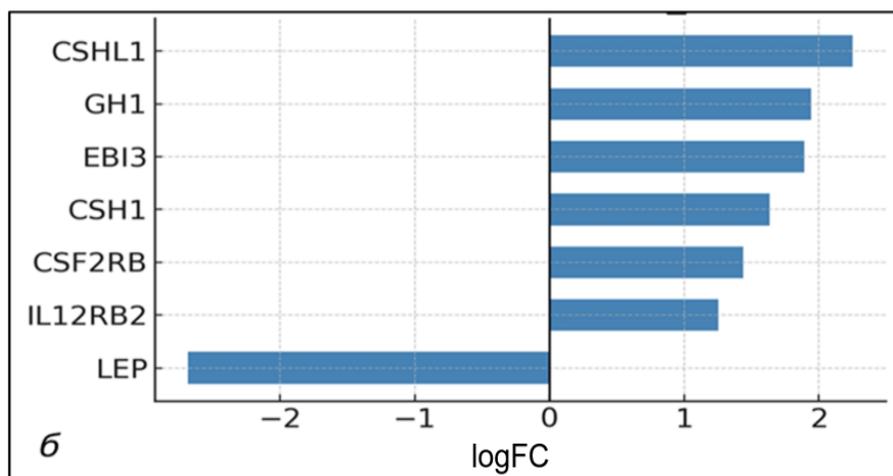


Рис. 3.14. Кількість генів для кожного з топ-4 функціонально збагачених біологічних процесів (а) і зміна експресії диференційно експресованих генів у другому триместрі у порівнянні з першим у кластері “Сигналінг через JAK-STAT шлях” (б)

Отже, з першого до другого триместру в плаценті активується імунна відповідь на плацентацию і розвиток ембріона, про що свідчить перелік збагачених

біологічних процесів (рис. 3.7, 3.8a) і активація прозапальних хемокінів (рис. 3.8б). Активуються гени класичних морфогенів, які забезпечують формування структури плаценти, диференціювання клітин і проліферацію (рис. 3.9 і 3.10). Цей період характеризується анаеробним гліколізом, підвищеним споживанням жирних кислот і глюкози на фоні материнської підвищеної інсулінорезистентності, тяжінням до оксидативного стресу через зниження захисних функцій металотіонеїнів і глутатіонпероксидази (рис. 3.12). Активація гормонів росту сприяє зростанню органу і активній взаємодії плаценти з організмом матері оскільки білки **GH1**, **CSH1** і **CSHL1** — секретованими (рис. 3.14). Метаболічні процеси характеризуються анаеробним гліколізом і підвищеним споживанням жирних кислот.

### **3.7. Кластери диференційно експресованих генів і функціонально збагачені біологічні процеси між другим і третім триместрами**

Період 2\_3 у порівнянні з 1\_2 інтервалом характеризується більшою кількістю кластерів диференційно експресованих генів. Функціонально збагачені біологічні процеси являють більшу різноманітність і часто іншу спрямованість одних і тих же процесів у порівнянні з першою половиною вагітності. Перелік диференційно експресованих генів наведений у додатку Г.

**3.7.1. Кластер 1. Регуляція відповіді на зовнішній стимул.** Цей кластер на відміну від попередньо розглянутих кластерів у порівнянні 1\_2 вирізняється більшою гетерогенністю. Про це наочно свідчить менша площа, яку займають процеси зі спільними генами (жовтуватий колір), і менша кількість спільних генів виявляється між процесами, позначеними близьким за інтенсивністю кольором (рис. 3.15).

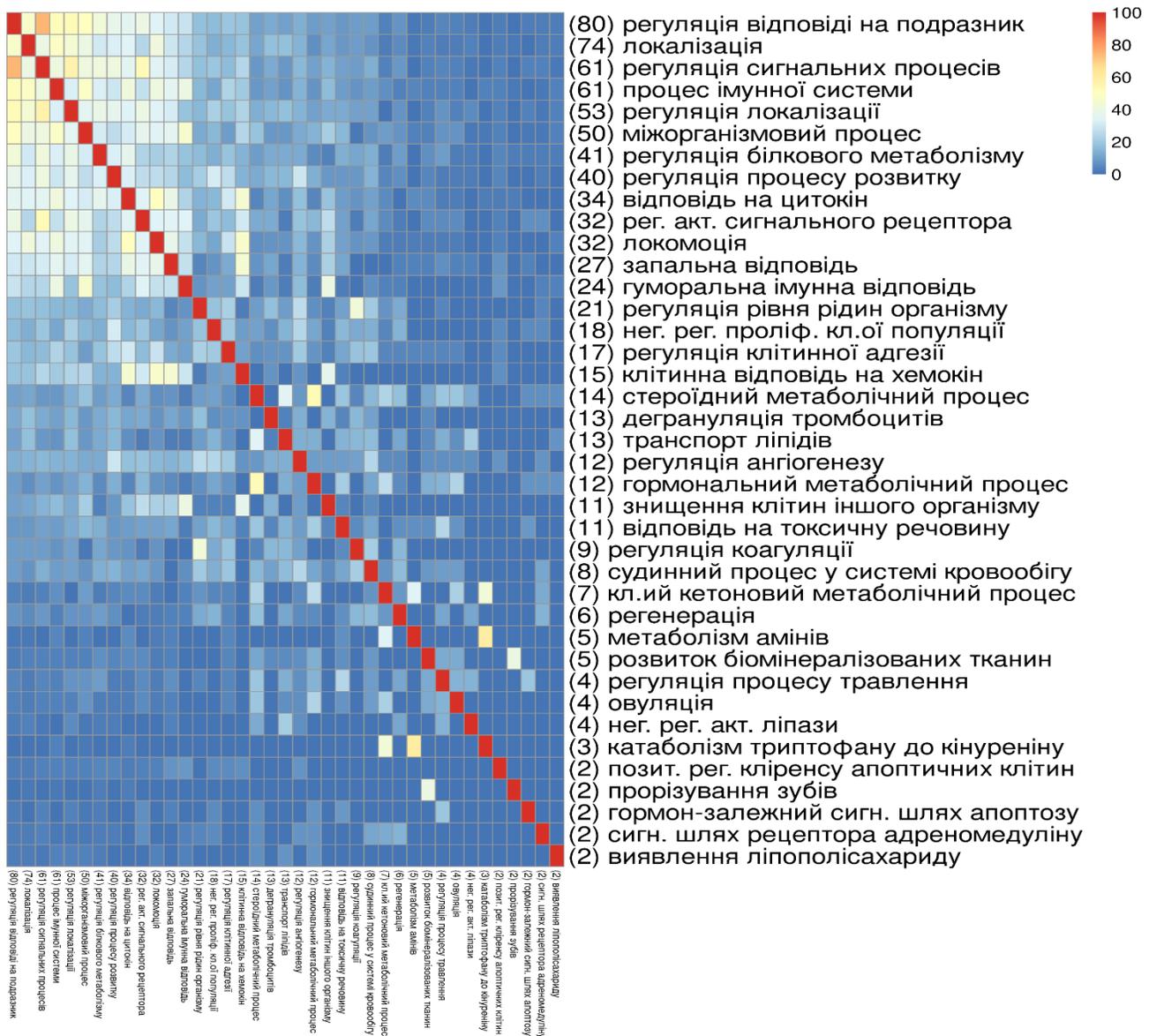


Рис. 3.15. Теплокарта функціонально збагачених біологічних процесів у кластері “Регуляція відповіді на зовнішній стимул”.  
Примітка: див. рис. 3.7.

Цей кластер об'єднує гени, які засвідчують карколомні зміни, яких зазнають ФЗБП і експресія диференційно експресованих генів. На відміну від першої половини вагітності знижується експресія хемокінів та інтерлейкінів (рис. 3.16б).

До цих змін додається нерівномірне зниження експресії генів системи комплементу (*C3*, -1,24; *C7*, 1,07; *CFD*, -1,05; *CFH*, 1,64).

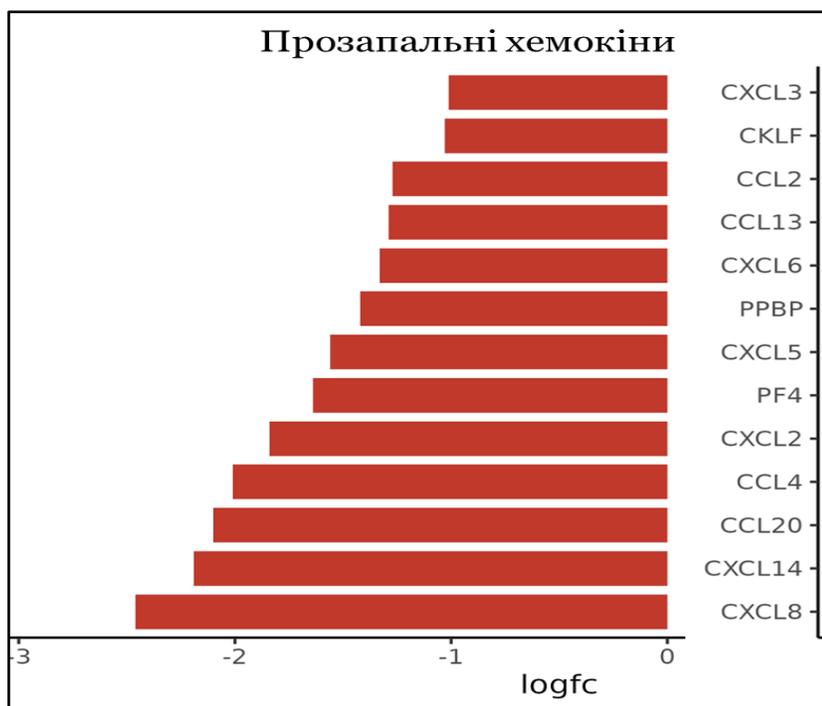
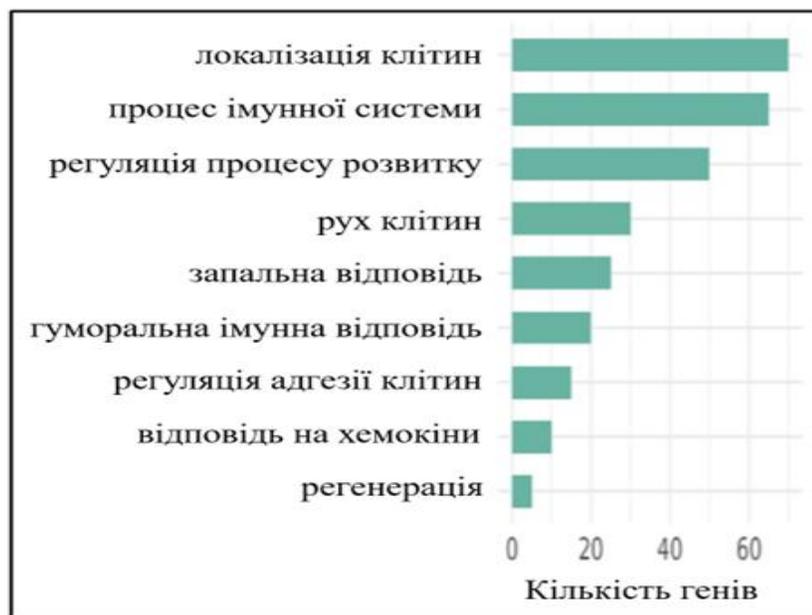


Рис. 3.16. Кількість генів для кожного з топ-9 функціонально збагачених біологічних процесів (а) та зміна експресії хемокинів між II і III триместром вагітності (б) у кластері “Регуляція відповіді на зовнішній стимул”.

Значних змін зазнає обмін ліпідів. Знижується експресія аполіпопротеїнів (*APOA2*, -1,63; *APOC1*, -1,56; *APOC3*, -1,07; *APOE*, -1,03; *APOH*, -1,21) на фоні подальшого підвищення експресії *FABP4* (5,45). Аполіпопротеїни — білки, які плацента синтезує і секретує в першій половині вагітності для забезпечення транспорту ліпідів з крові матері в плаценту. У другій половині цю функцію переймає плод, *FABP4* — це внутрішньоклітинний білок, який зв'язує та транспортує жирні кислоти всередині клітин. Він спрямовує їх до ядерного

гормонального рецептора (PPAR $\gamma$ ), який є транскрипційним фактором і регулює різні клітинні процеси, включаючи метаболізм ліпідів та глюкози, диференціацію клітин та імунітет. Суттєво зростає експресія гена **LEP (2,58)** від II до III триместру. Водночас ймовірно посилюється обмін холестерину в самій плаценті через активацію ключового регулятора гомеостазу холестерину гена **SCARB1** (Scavenger Receptor Class B Type 1, **1,8**), який відповідає за перенесення ефірів холестерину між ліпопротеїнами, і гена **CETP** (Cholesteryl Ester Transfer Protein, **2,16**), який задіяний у захопленні й перенесенні холестерину в середину клітини. Обмін холестерину є критично важливим для синтезу стероїдів у самій плаценті і для розвитку мозку плода.

Стрімке зростання експресії гена **CRH** (Corticotropin-Releasing Hormone gene, **6,14**) і підвищення концентрації кортикотропіну в крові матері активує гіпоталамо-гіпофізарно-наднирникову вісь, сприяє підготовці організму матері до пологів. Цей гормон важливий для розвитку нервової системи плоду [74].

Змінюється структура гемоглобіну, який утворюється плацентою. Знижується експресія субодиниці епсилон, притаманної фетальному гемоглобіну (**HBE1, -3,73**), і підвищується експресія субодиниці дельта (**HBD, 1,29**), яка входить до складу мінорної форми гемоглобіну A2 дорослої людини.

**3.7.2. Кластер 2. Метаболічні процеси і клітинний цикл.** Теплокарта (рис. 3.17) демонструє функціонально збагачені біологічні процеси кластеру. Експресія генів, прямо чи опосередковано причетних до фаз клітинного циклу, достовірно в 2 – 4 рази зменшується між II і III триместром за винятком гена **KIF2A (1,07)** (рис. 3.18). Білок **KIF2A** необхідний для нормального шикування хромосом у метафазній пластинці і функціонування веретена поділу під час мітозу. Пригнічення експресії генів клітинного циклу підтверджує загальне суттєве зниження проліферації клітин у плаценті в міру прогресування гестації [75].

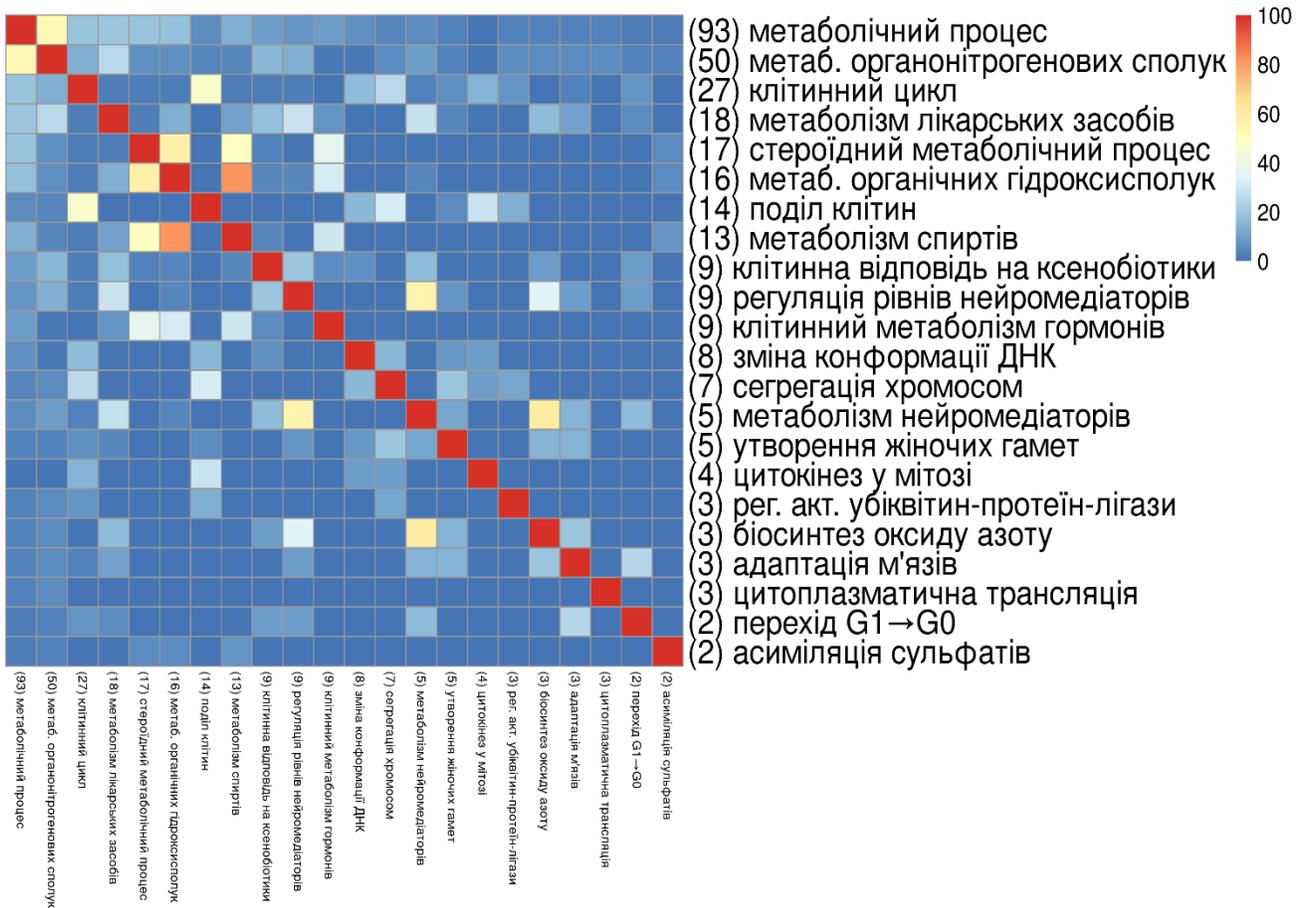


Рис. 3.17. Теплокарта функціонально збагачених біологічних процесів у кластері “Метаболічні процеси і клітинний цикл”

Примітка. Див. Рис. 3.7.

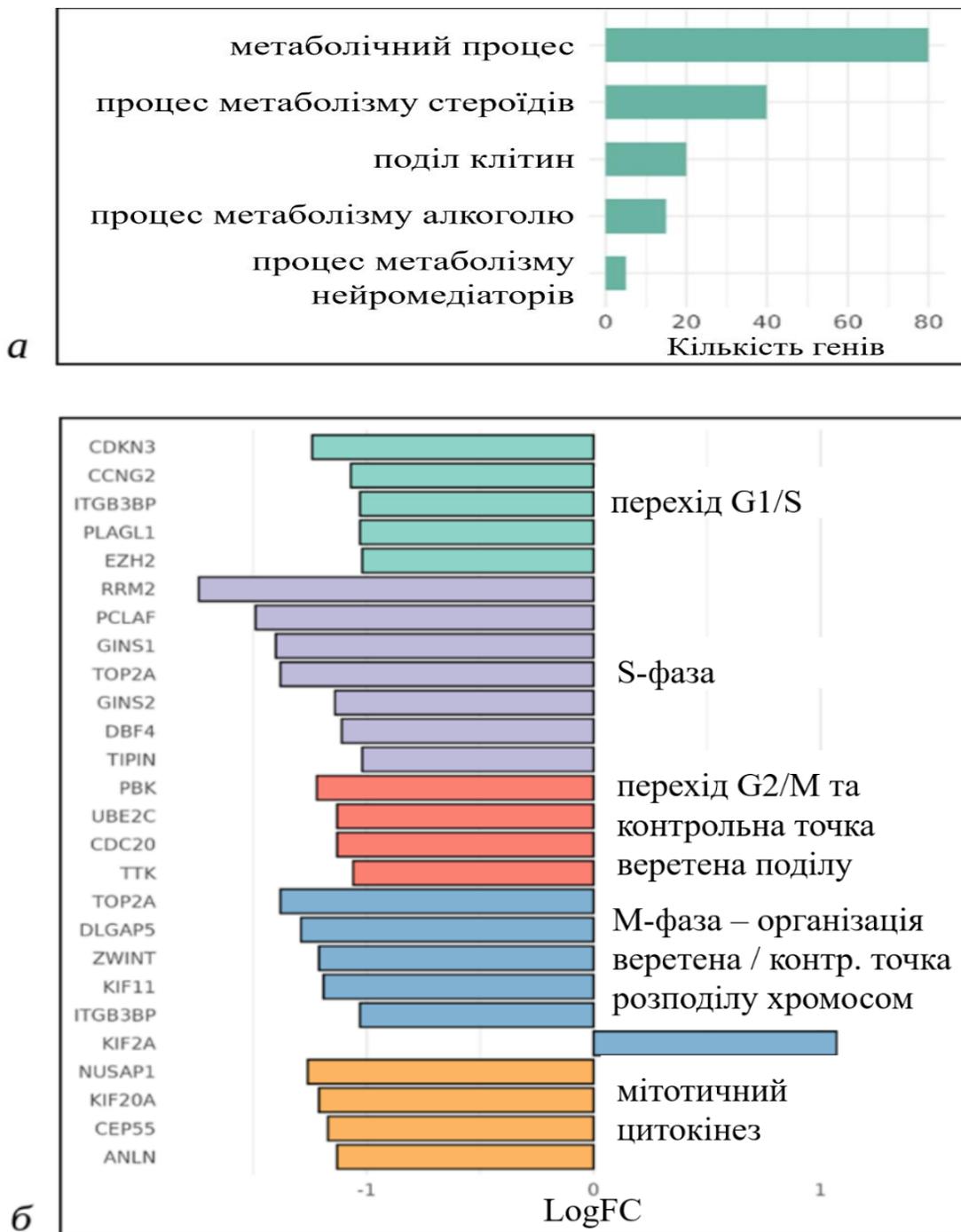


Рис. 3.18. Топ-5 функціонально збагачених біологічних процесів (а) та зміна експресії генів клітинного циклу між II і III триместром вагітності (б) у кластері “Метаболічні процеси і клітинний цикл”

На цьому етапі окрім зниженої експресії генів клітинного циклу вирізняється активація генів біосинтезу стероїдів (*FDX1*, 1,23; *CYP11A1*, 2,02; *INHBA*, 3,58; *HSD17B1*, 1,22; *HSD3B1*, 2,28; *CYP19A1*, 2,42; *HSD3B2*, -1,71), а експресія генів синтезу холестерину (*CYP51A1*, -1,14; *MSMO1*, -1,67; *HMGCS1*, -1,03 і *DHCR24*, -1,24) знижена. Якщо в першій половині плацента сама забезпечує потреби свої і плоду у холестерині, то у другій вона активно транспортує материнський холестерин (див. пункт 3.7.1), який є в достатку [76]. Дані щодо активації експресії генів стероїдогенезу наприкінці вагітності підтверджують раніше отримані методом ЗТ-ПЛР результати [77].

**3.7.3. Кластер 3. Процесинг О-гліканів.** Кластер містить гени, які виконують декілька біологічних функцій: формування і секреція муцинів та О-глікозилювання (*MUC16*, *MUC20*, *GALNT3*, *GALNT15*, *ST3GAL4*, *AGR2*), клітинна адгезія (*CLDN8*, *NECTIN3*, *NECTIN4*, *ESAM*, *LAD1*, *EMP2*) та деякі функції, пов'язані з імунітетом (*S100P*, *IFI27*, *TRIM40*, *TRIM62*) (повний перелік ДЕГів на рис. 3.19). Лише група генів О-глікозилювання є функціонально збагаченою. Муцини — це великі, сильно глікозильовані білки, які є основними компонентами слизу, що вкриває і захищає епітелій дихальних, травних, сечостатевого та інших слизових оболонок. О-глікозилювання — ключовий процес у формуванні муцинів.

Два найбільш диференційно експресованих гени кластера є маркерами позаворсинового трофобласту *HLA-G* (-2,13) — нетиповий (non-classical) представник МНС класу I, сприяє імунній толерантності через зв'язування з інгібіторними рецепторами імунних клітин матері, і *ALPP* (6,65) — глікозильований алостеричний фермент із родини лужних фосфатаз, який локалізується на зовнішньоклітинній мембрані синцитіотрофобласту й сприяє транспорту поживних речовин через матково-плацентарний бар'єр до плоду [78,

79]. Такі зміни експресії цих двох генів відомі, але надмірність змін супроводжується ускладненнями.

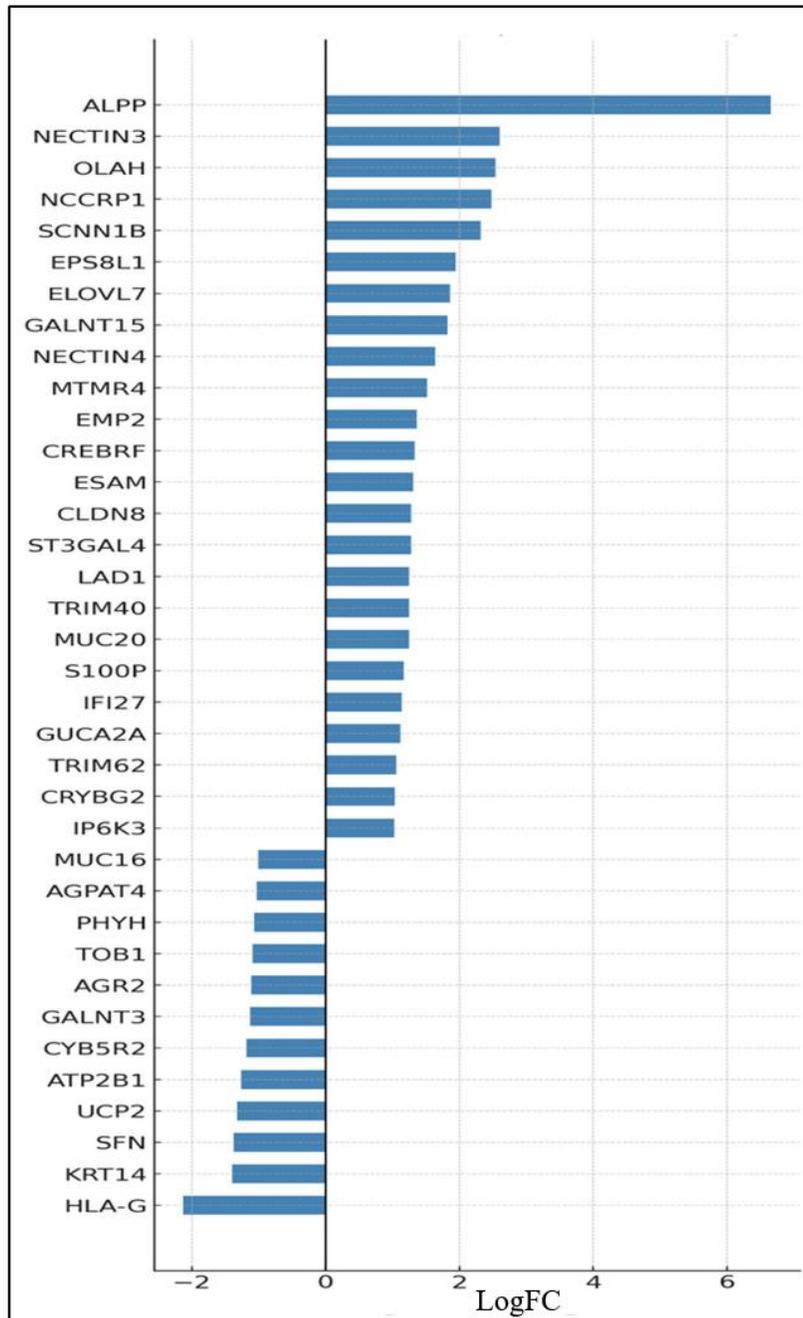


Рис. 3.19. Диференційна експресія генів у кластері “Процесинг О-гліканів” у третьому триместрі у порівнянні з другим

Якщо порахувати на графі білок-білкових взаємодій для кожного гена з кластера 3 періоду 2\_3 кількість взаємодій з усіма іншими генами й обчислити середнє, воно складатиме 10.7 генів у той час, як середня кількість взаємодій для генів у кластерах 1, 2 і з 4 по 7 — 22.0. Середня кількість взаємодій всередині кластера 3 ще менша — 2.39. Тож можемо стверджувати, що кластер 3 складається переважно з генів, що мають низький рівень взаємодії як між собою, так і з іншими елементами мережі. Саме ця низька взаємодія могла стати причиною об'єднання алгоритмом кластеризації fast greedy низки біологічно розрізнених генів в окремий кластер.

**3.7.4. Кластер 4. Морфогенез тканин і фактори комітування.** За даними Stringdb ФЗБП визначаються як такі, які відповідають за процеси ембріогенезу, формування різних органів і структур (рис. 3.20, 3.21), а за базою даних Reactome додаються категорії формування (27) і деградації (13) позаклітинного матриксу, синтезу і деградації колагену та утворення мультимерних структур (8), передачі сигналів за допомогою тромбоцитарного фактору росту (7), негативної регуляції TCF-залежного сигналіngu антагоністами ліганду WNT (4) і активації матриксних металопротеїназ (4). У дужках наведено кількість генів у категорії.

Серед ДЕГів спостерігаємо знижену експресію морфогенів (*WNT5A* і *WNT2*), антагоністів (*DKK1* і *DKK2*) і розчинного модулятора (*SFRP1*) WNT сигналіngu (рис. 3.21) *WNT5A* пов'язаний з неканонічною сигналізацією Wnt і активує шляхи Wnt-Ca<sup>2+</sup> та Wnt-PCP (Wnt/Planar Cell Polarity), підтримуючи полярність клітин в площині, ортогональній до апікально-базальної вісі. *WNT2*, ліганд канонічного шляху, зв'язується з рецепторами на поверхні клітин, передає інформацію всередину клітин, стабілізує β-катенін, який проникає у ядро і через транскрипцію відповідних генів активує проліферацію і підтримує стабільність стовбурових клітин, регулює диференціювання, міграцію і адгезію клітин,



Зниження експресії охоплює і гени-представники надродини TGF- $\beta$ , вкрай важливої для успішної вагітності [84]. **BMP2** задіяний у інвазії трофобласта та ремоделюванні спіральних артерій; порушення регуляції **BMP2** у плаценті може сприяти розвиткові ранньої форми преєклампсії [84]. **BMP6** опосередковує дію **BMP2** [85]. **Інгібін А (INHA)** і **інгібін В (INHBA)** відомі своєю здатністю інгібувати дію активінів, схожих за природою з інгібінами, але на відміну від інгібнів вони спрямовують стовбурові клітини до диференціювання, стимулюють синтез колагену та утворення фіброblastів. Інгібін альфа і бета експресуються у позаворсинковому хоріоні, синцитіотрофобласті й в ендометрії, яєчниках, наднирниках, мозку і яєчках і є пептидними гормонами [86]. Концентрація INHBA і INHA у материнській плазмі і плаценті підвищується впродовж гестації [87, 88, 89]. Визначення рівня інгібіну А в крові матері протягом першого стандартного обстеження вагітних на 11-13 тижні + 6 днів вагітності допомагає виявити синдром Дауна і ризик викиднів, преєклампсії та затримки росту плода до появи клінічних симптомів [90, 91, 92]. До надродини TGF- $\beta$  належить і гормон **GDF15** (Growth and Differentiation Factor 15). Білок секретується різними клітинами організму, особливо у великій кількості плацентою і плодом, і рівень гормону в крові матері наприкінці вагітності може в 200 разів перевищувати рівень, притаманний невагітним. На початку вагітності з підвищеним рівнем **GDF15** пов'язують нудоту, блювоту, відразу до їжі і тяжку форми цих проявів, гіперемезис вагітних. Легка форма таких проявів зазвичай закінчується після першого триместру [93, 94, 103]. Яким чином діє **GDF15** в кінці вагітності — відкрите питання.

Ще одну групу складають транскрипційні фактори, які кодуєть фактори комітування, які спрямовують клітини до спеціалізації. **GATA6** відомий своєю участю в розвитку серця та підшлункової залози і у ранньому позаембріональному

розвитку тканин [95]; **HAND1** відіграє певну роль у розвитку серця, похідних нервового гребня та гігантських клітин трофобласта [96, 97, 98]; **ISL1** із родини білків з LIM-гомеодоменом (LIM-домен для зв'язку з білками і гомеодомен для зв'язку з ДНК) задіяний у регуляції розвитку різних органів і систем, включаючи серце, підшлункову залозу і нервову систему. **ISL1** є маркером позаембріональної мезодерми. Він необхідний для формування алантоїсу/пуповини, хоріон-алантоїсної пластинки плаценти і плаценти як такої та її судинної системи [99]. **PITX2** (Paired-like homeodomain transcription factor 2) у ранньому ембріогенезі задіяний у встановленні ліво-правої осі, асиметричному лівосторонньому розвитку серця та селезінки, повороту кишківника та шлунка і симетричному розвитку очей і формуванні структур у передній частині ока, включаючи райдужку, кришталік та рогівку. **PITX2** також бере участь у розвитку гіпофіза і зубів. Серед цих численних функцій вирізняється його критична необхідність для імплантації бластоцисти [100]. **PITX2** взаємодіє із сигнальним шляхом Wnt/ $\beta$ -катеніну.

Зниження експресії всіх перелічених генів, які задіяні у морфогенезі, проліферації і комітуванні, вказує на те, що їх активна роль протягом початкової стадії гестації виконана, і в подальшому вони непотрібні.

До кластера “Морфогенез тканини і фактори комітування” потрапили також гени, які відповідають за синтез колагену (**COL21A1**, -1,50; **COL6A2**, - 1,01; **COL5A1**, -1,2; **COL14A1**, -1,07; **COL6A6**, -1,38) і синтез гіалуронової кислоти (**HAS2**, -1,48). Їх знижена експресія зумовлює зменшення щільності структури плаценти, що необхідно перед майбутніми пологами [101]. До цього процесу долучаються і матриксні металопротеази [102].

Морфогенетичні фактори та фактори комітування (**TGF**, **Activins**, **BMPs**, і **GDFs**) реалізують свою активність через серин/треонін-киназні рецептори,

специфічний клас рецепторів, які передають сигнали всередину клітини за рахунок фосфорилування серинових і треонінових залишків у білках. Гени, які кодують такі рецептори, не потрапили до складу ДЕГів.

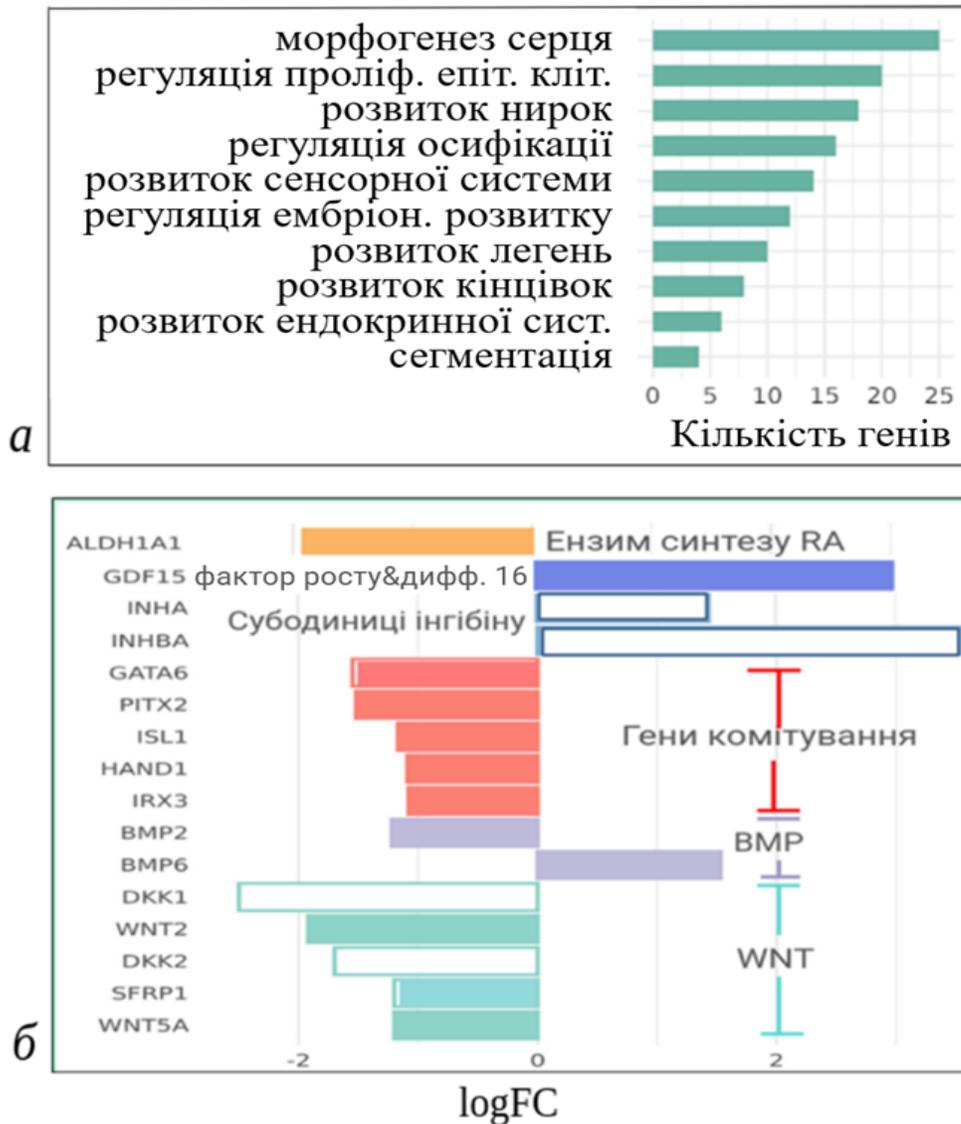


Рис. 3.21. Топ-5 функціонально збагачених біологічних процесів (А) та експресія генів, пов'язаних з морфогенезом і комітуванням, у третьому триместрі у порівнянні з другим (Б).

Примітка: гени-активатори позначені зафарбованими блоками; антагоністи Wnt сигналіну, *DKK1* і *DKK2*, і гени *ІНА* і *ІНВА* білків-інгібіторів позначені незафарбованими блоками.

**3.7.5. Кластер 5 “Транспорт натрію” і кластер 6 “Трансмембранний транспорт хлориду”.** Гени кластера 5 (рис. 3.22а) належать до двох функціональних категорій — запасного шляху синтезу нуклеозидів та трансмембранного транспорту іонів натрію. Запасний шлях синтезу нуклеозидів, на відміну від *de novo* синтезу, є характерним для диференційованих клітин із низькою проліферативною активністю, що спостерігається у плаценті наприкінці вагітності [104]. Трансмембранний транспорт хлориду й загальна категорія «транспорт» входять до 6-го кластера (рис. 3.22б). У плаценті іони хлориду виконують насамперед дві важливі функції: по-перше, вони створюють осмотичний градієнт, що призводить до руху води через синцитіотрофобласт в амніотичний і матковий простір [105]; по-друге, через хлоридні канали регулюється мембранний потенціал клітин, що впливає на секрецію гормонів і транспорт інших іонів та молекул [106]. На рис. 3.23. наведено диференційно експресовані гени в кластерах 5 і 6.

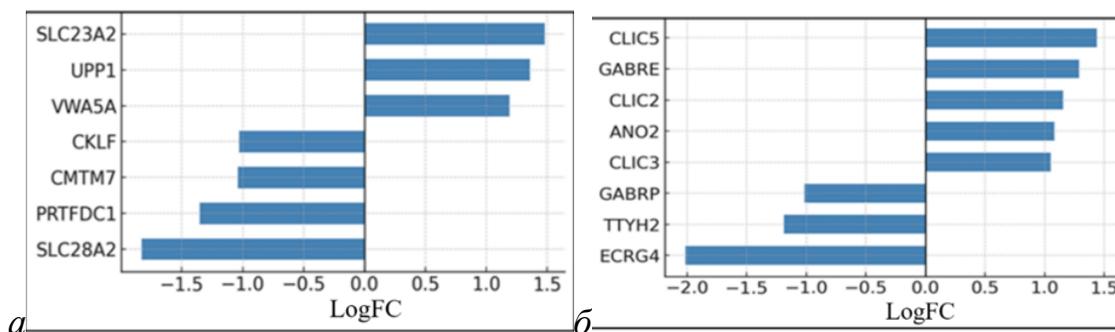


Рис. 3.22. Диференційна експресія генів у кластері 5 “Транспорт натрію” (а) і кластері 6 “Трансмембранний транспорт хлориду” (б) у третьому триместрі у порівнянні з другим.

Білок **SLC23A2** активно транспортує аскорбінову кислоту в клітини трофобласту і діє як антиоксидант і захисник колагену. Ген активно експресується

в плаценті, мозку і кишківнику. Білок **SLC28A2** -  $\text{Na}^+$ -залежний селективний переносник нуклеозидів і аденозину.

### 3.7.6. Кластер 7. Специфічні для вагітності бета-1-глікопротеїни.

Специфічні для вагітності глікопротеїни (pregnancy-specific glycoproteins, PSG) — це білки з ковалентно зв'язаними олігосахаридами, приєднаними до бокових ланцюгів амінокислот. Глікопротеїни є мажорними білками в кровотоці матері на пізніх термінах вагітності. Десять генів *PSG* кодують білки (PSG1 — PSG9, PSG11). Відомо, що ці білки мають імунорегуляторну, проангіогенну та антитромбоцитарну функції. Вони взаємодіють і активують протизапальні цитокіни, фактори росту  $\text{TGF}\beta 1$  та  $\text{TGF}\beta 2$ , що робить їх одними з небагатьох відомих біологічних активаторів цих важливих цитокінів. Між другим і третім триместром суттєво (у 8 — 32 рази) підвищується експресія генів *PSG* 1, 2, 3, 4, 9 і 11 (рис. 3.23).

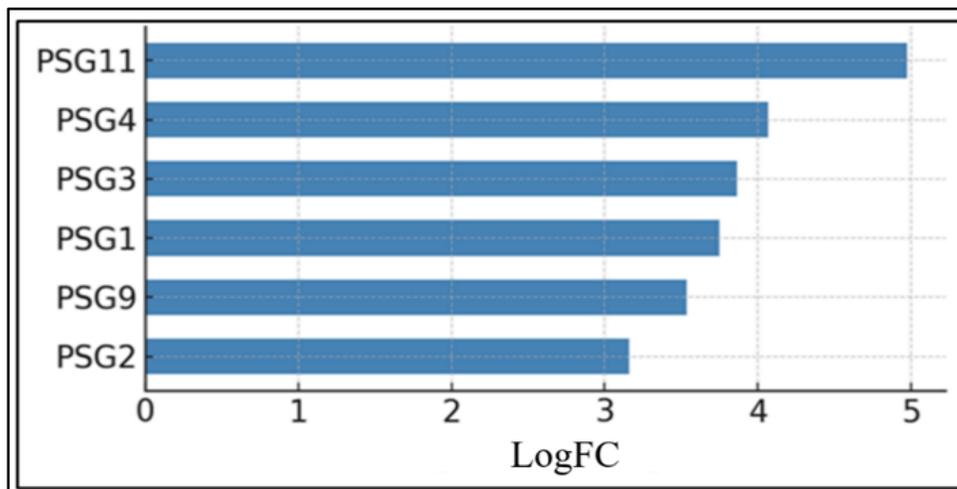


Рис. 3.23. Експресія генів, які кодують специфічні для вагітності бета-1-глікопротеїнів (кластер 7), в третьому триместрі у порівнянні з другим

### 3.8. Порівняльний аналіз генної експресії у плаценті першої і другої половини вагітності

На тлі вище описаної різниці між генною експресією в першій і другій половині вагітності виокремлюються 54 гени, які є спільними для обох строків (рис. 3.24). Аналіз інтенсивності експресії цих генів, наданих в одних і тих же одиницях (через нормалізацію двох груп диференційно експресованих генів) дає можливість виявити додаткові тенденції у зміні генної експресії впродовж вагітності і більш наочно їх представити. Серед 54х генів вирізняються чотири групи: "UpUp" — експресія ДЕГів (18) підвищується впродовж всього періоду гестації; "UpDown" (21 ген) — підвищення експресії впродовж 1\_2 і зниження впродовж 2\_3 періоду; "DownDown" — експресія ДЕГів (12) знижується впродовж усього періоду гестації і "DownUp" (3 гени) — зниження впродовж 1\_2 періоду і підвищення впродовж 2\_3 періоду (Таблиці 3.4 — 3.6).

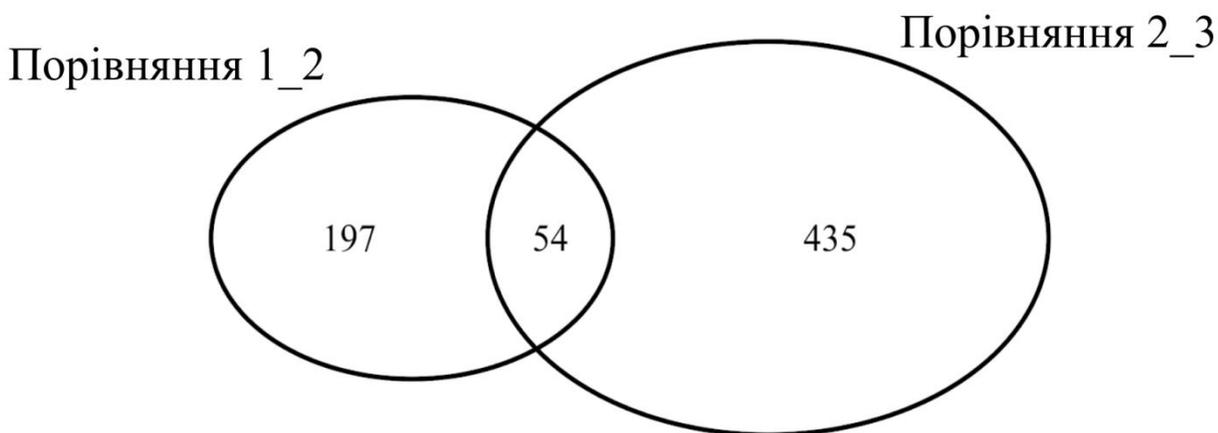


Рис. 3.24. Діаграма Венна: спільні диференційно експресовані гени з  $|\log_{FC}| > 1$  у порівняннях 1\_2 і 2\_3

Група генів "UpUp" (рис. 3.25), Як можна було очікувати, в цю групу потрапили гени з різними функціями, тим не менше деякі з них все ж можна об'єднати за специфікою їх природної активності. Серед них гени, причетні до:

здатності видаляти бактерії і білкові комплекси (*MAP1LC3C*, *C7*); адгезії клітин (*ESAM*, *NECTIN3B*, *ADGR6G*), останній відомий також як регулятор розвитку нервової системи, серця і вух; ембріонального розвитку (*MAGEA10*); системи згортання крові, негативної регуляції розщеплення фібрину (*SERPINB2*, *CLEC3B*); метаболічних процесів і зниження утилізації глюкози на користь підвищення метаболізму жирів (*PDK4*) і активації стадії в циклі Кребса на користь синтезу амінокислот (*ADHEF1*); регуляції синтезу ростових факторів (*PAPPA*), цілеспрямованого росту клітинних структур (*SEMA3B*) і проліферації Т-клітин (*CD200*); різнопланової активності цитокінів ІІ-27, ІІ-35 і ІІ-39, для яких **EBI3** є спільною субодиницею; гормональної активності, яку забезпечують гени *CSHL1* і *HSD17B1* і до дилатації судин через активацію пептида **ADM** і гена *CHI*, який сприяє виділенню NO.

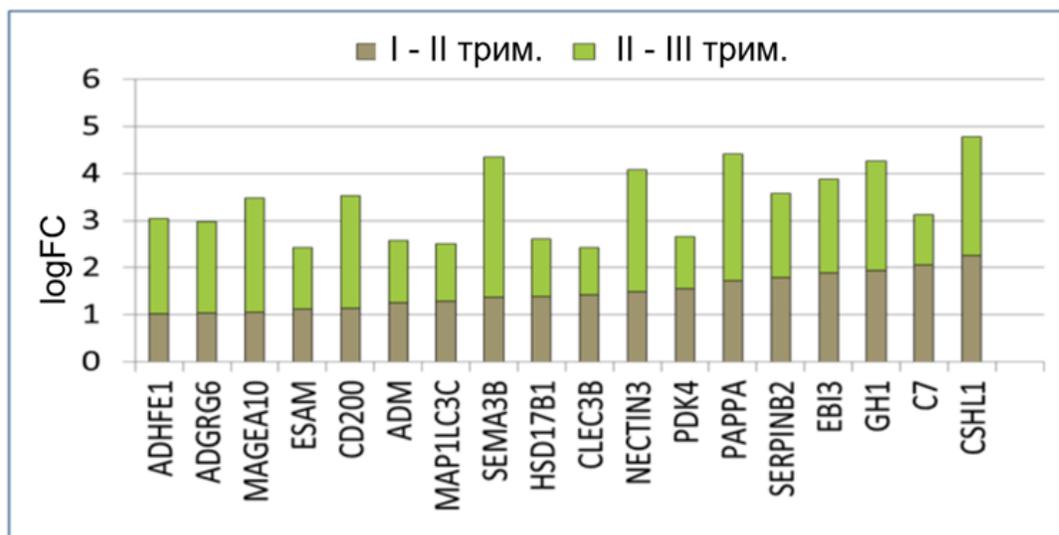


Рис. 3.25. Експресія спільних диференційно експресованих генів з групи UpUp у першій і другій половині вагітності

Група генів "UpDown" (рис. 3.26). Одинадцять з двадцяти одного гена цієї групи так чи інакше пов'язані з прозапальною активністю. До них належать всі хемокіни (*CCL2*, *CCL13*, *CXCL2*, *CXCL4*, *CXCL7*, *CXCL8*, *PF4*); ІІ6 разом з

**CXCL8** (ака **И8**) є найпотужнішими прозапальними цитокінами [124]; **NTS** залучений до місцевих прозапальних процесів, активації проліферації і міграції клітин [107], а **HPGDS** — до синтезу простагландинів, прозапальних ліпідних медіаторів [108]. **CHN1** кодує прозапальний хімерин з активністю хемокіна.

Дія **TIMD4** і **BCL2A1** пов'язана з апоптотичим процесом, **TIMD4** сприяє очищенню тканини від апоптотичих клітин шляхом фагоцитозу, а **BCL2A1** зменшує вивільнення проапоптотичного цитохрому С із мітохондрій, блокує активацію каспаз і протистоїть запрограмованій смерті клітин [109].

**AOC1** і **PRG2** впливають на тонус судин: **AOC1** є єдиним ензимом, який розщеплює гістамін. **PRG2** утворює в плаценті і крові комплекс з **PAPPA** і ангіотензиногеном (**AGT**) і таким чином знижує їх активність. Це призводить до інактивації ангіотензиногену в першій половині вагітності і навпаки до його активації у другій половині з вірогідним перетворенням ангіотензиногену в ангіотензин, який підвищує тиск. Вивільнення **PAPPA** із комплексу з **PRG2** впродовж періоду 2\_3 сприятиме руйнування білку **IGFBP** і вивільненню IGF.

На діаграмі (рис. 3.26б) на прикладі експресії гена **CXCL8** помічені всі етапи зміни експресії, притаманні значній кількості генів цієї групи. Експресія гена **CXCL8** підвищується від першого до другого триместру в 3.3 рази ( $\log_{FC} 1,74$ ) і знижується у третьому семестрі у порівнянні з другим в 5.5 разів ( $\log_{FC} 2,46$ ), залишаючись у третьому триместрі в 1.65 разів ( $\log_{FC} -0,72$ ) нижчою за початковий рівень у першому триместрі. Експресія значної кількості генів цієї групи в результаті зниження при переході від другого до третього триместру залишається вищою за експресію у першому триместрі. Якби ми порівнювали перший з третім триместром, то констатували би або зниження експресії на зразок гена **CXCL8** або підвищення на зразок інших генів цієї групи, пропускаючи принципову особливість, яка об'єднує гени цієї групи, а саме зміну направленості

експресії у другому триместрі при переході до третього.

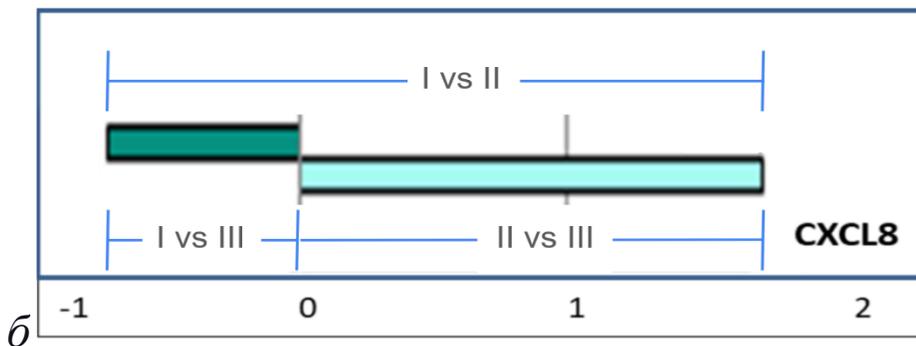
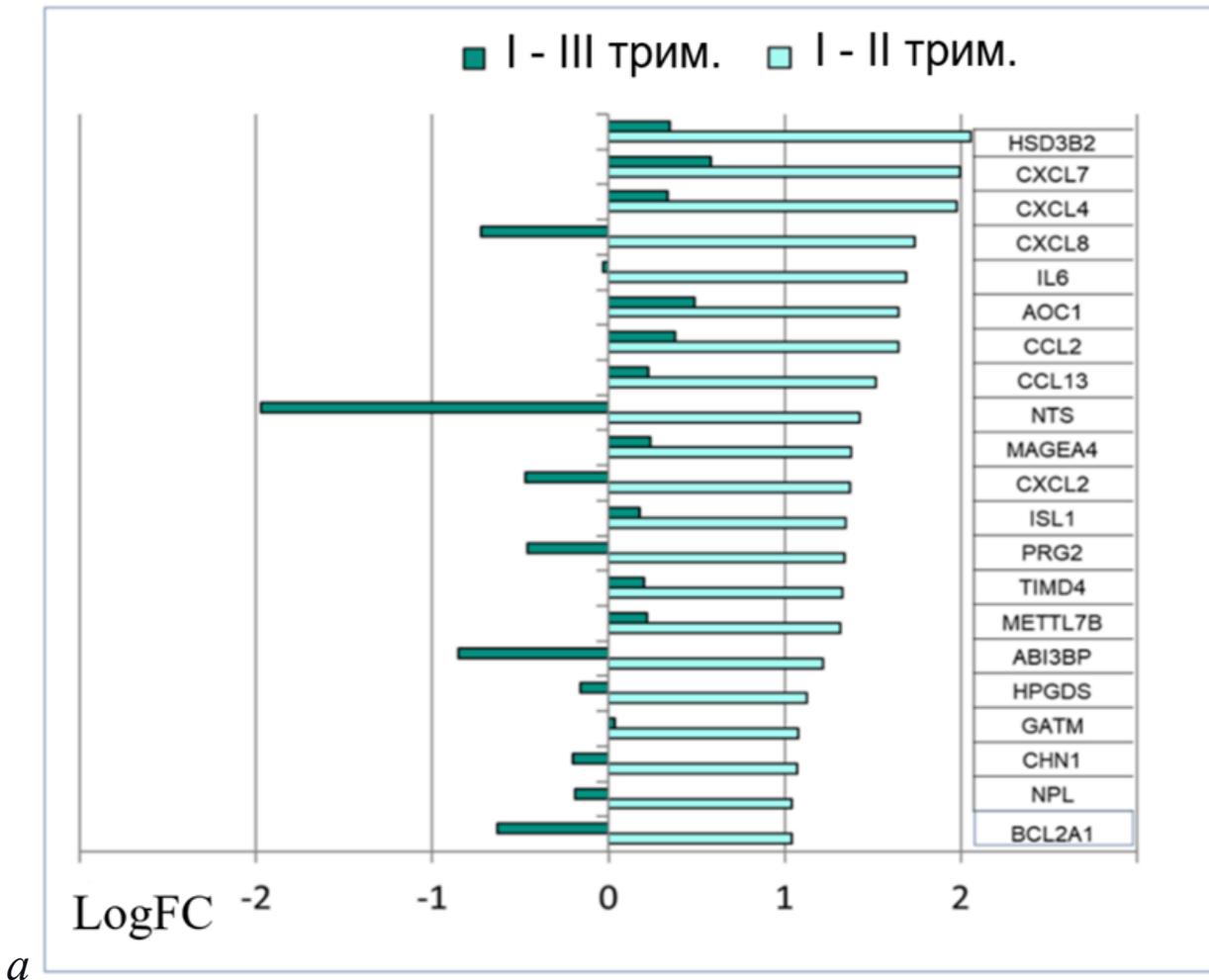


Рис. 3.26. Експресія диференційно експресованих генів з групи UpDown у першій і другій половині вагітності (*a*) та детальніше пояснення стовпчиків діаграми (*б*)

Інші гени цієї групи задіяні у: синтезі естрогенних гормонів і тестостерону (**HSD3B2**) [110] і креатину (**GATM**); регуляції вмісту сіалової кислоти на поверхні глікокаліксу (**NPL**); метилюванні білків і РНК (**METTL7B**), і у процесах морфогенезу і розвитку (**ISL-1**, **ABI3BP**). Найбільший діапазон змін виявляє експресія гена **NTS**. Цей ген кодує спільний попередник для двох пептидів: нейромедину N та нейротензину. Нейротензин — це секретований тридекапептид, який широко поширений у центральній нервовій системі та може функціонувати як нейромедіатор або нейромодулятор. Нейротензин також проявляє антимікробну і антифунгіцидну активність, Тропізм нейротензину до нервової системи робить його потенційним об'єктом дослідження механізмів розвитку (пре)еклампсії.

*Група генів "DownDown"* (рис. 3.27). До цієї групи належать гени, які відіграють суттєву роль на більш ранніх етапах формування плаценти та імплантації бластоцисти. Такими є ген **PAEP**, який модулює ендокринну функцію та диференціювання клітин трофобласта і регулює по різних напрямках активність імунних клітин, щоб забезпечити взаємну толерантність матері й плоду [127].

Альфа-фетопротеїн (**AFP**) виконує функції ембріонального альбуміна, а епсілон субодиниця ( $\epsilon$ ) (**HBE1**) є субодиницею фетального гемоглобіну.

Гени **IRX3** (*Iroquois Homeobox 3*) і **OSR2** кодують транскрипційні фактори, які беруть широку участь у ембріональному морфогенезі, зокрема у оптимізації взаємодії заплідненої яйцеклітини з децидуа під час імплантації (**IRX3**) [111] і у формуванні щелеп і піднебіння, зубів, верхніх і нижніх кінцівок, що продемонстровано на мишах [112].

Експресія трьох генів, **LOX**, **FLRT2** і **AGR2**, пов'язана з функціонуванням позаклітинного матрикса і міжклітинною взаємодією, підвищує рихлість матриксу і сприяє міграції клітин і ангиогенезу. Ген **LOX** (*Lysyl Oxidase*) кодує  $\text{Cu}^{2+}$ -

залежний ензим, який каталізує перехресне зшивання колагену і еластину, підсилює міцність функціонально зрілого ПЗКМ і впливає на міграцію клітин, ангіогенез і на міжклітинні взаємодії [113]; *FLRT2* регулює міжклітинні взаємодії у тканинах, що формуються, і *AGR2* може сприяти епітеліально-мезенхімальному переходу (EMT), що полегшує міграцію клітин.

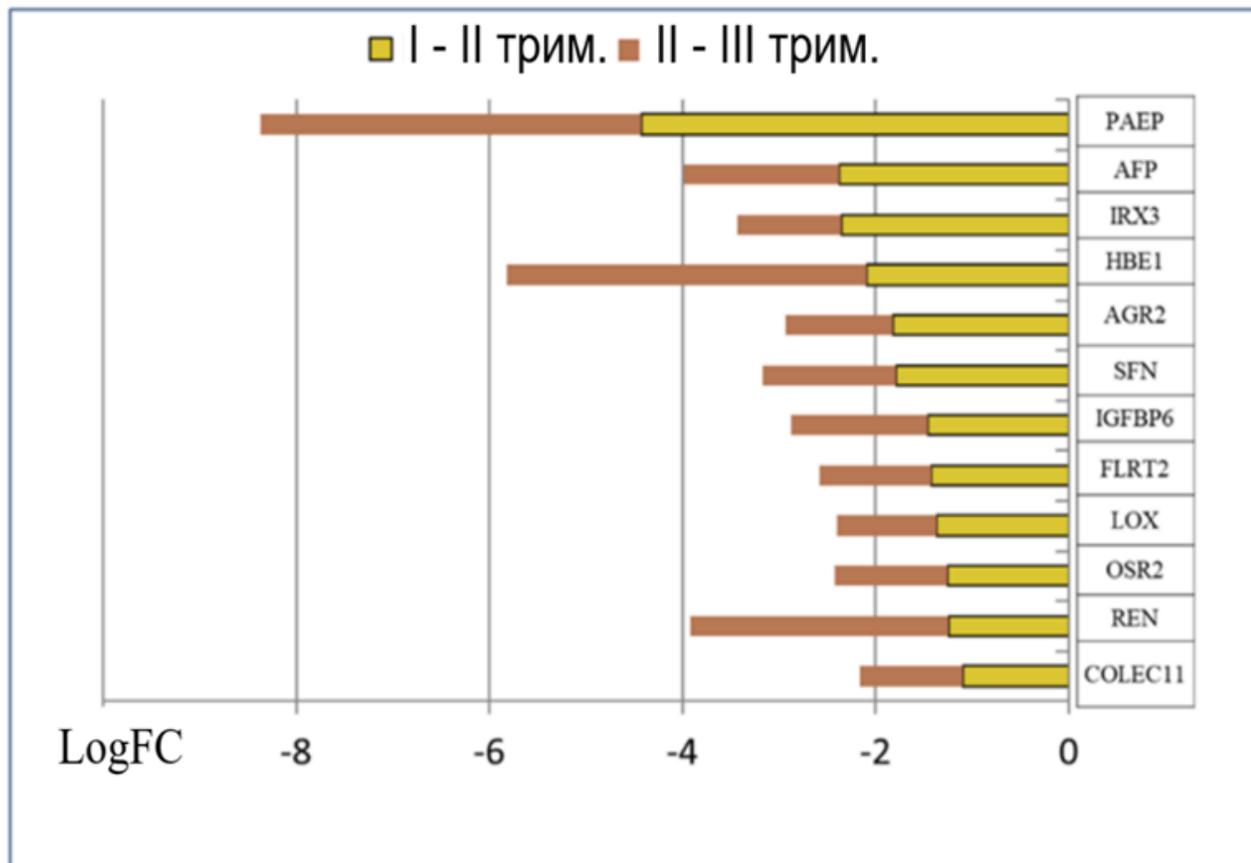


Рис. 3.27. Експресія диференційно експресованих генів з групи DownDown у першій і другій половині вагітності

Ген *REN* кодує високоспецифічну ендопептидазу, єдиною відомою функцією якої є перетворення ангіотензиногену в ангіотензин, який у свою чергу ініціює каскад реакцій, що призводять до підвищення артеріального тиску та збільшення затримки натрію нирками.

Ген *SEN* кодує адаптерний білок 14-3-3 Sigma, який контролює клітинний цикл, і ген *Collectin-11* кодує білок з родини лектинів С-типу, які мають колагеноподібні послідовності та домени розпізнавання вуглеводів. Колектини — це секретовані білки, які відіграють важливу роль у вродженому імунітеті, зв'язуючись із вуглеводними антигенами мікроорганізмів, полегшуючи їх розпізнавання та видалення.

Група *DownUp* (рис. 3.28) містить лише три гени: *LEP* (лептин) — гормон, що регулює енергетичний баланс, керує метаболічними реакціями і відчуттям ситості/голоду; *MUC20* (муцин 20) — білок, що забезпечує захист та змащування поверхні клітин, бере участь у сигналізації між клітинами, і тканинний активатор плазміногену *PLAT* — фермент, який розщеплює кров'яні згустки і впливає на "ремонт" та ремоделювання тканин.

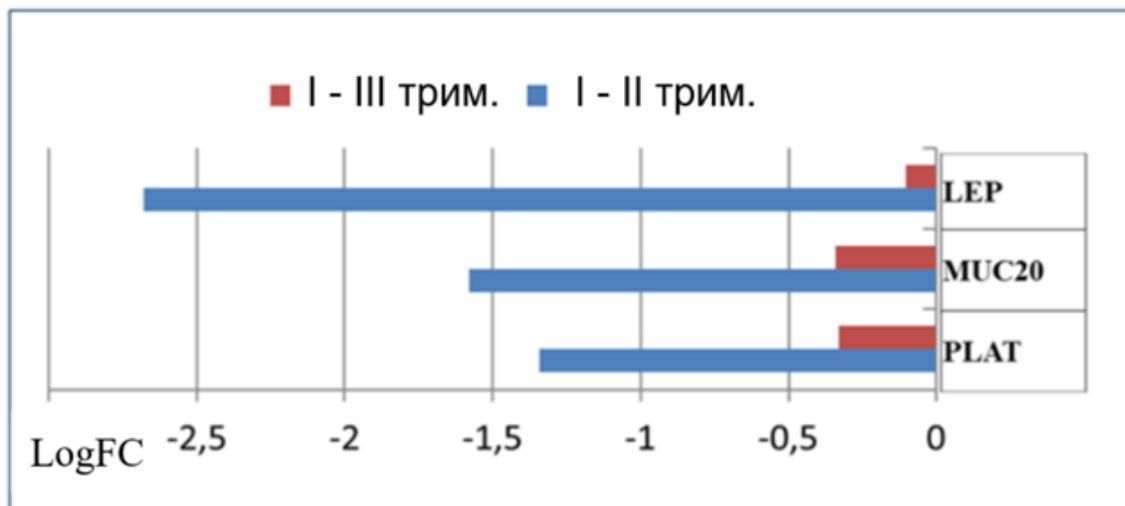


Рис. 3.28. Експресія диференційно експресованих генів з групи *DownUp* першій і другій половині вагітності

Рис. 3.28 наочно демонструє як різко змінюється направленість експресії генів від зниження впродовж першої половини вагітності до підвищення у третьому триместрі, в 5,6 разів ( $\log_{FC} 2,5$ ) на прикладі гена *LEP*, хоча залишається

нижчою за значення у першому триместрі. Якби ми порівнювали експресію в першому триместрі з третім, то ми б констатували незначне зниження експресії в третьому триместрі у порівнянні з першим, пропускаючи суттєві зміни при переході до другої половини вагітності, які об'єднують ці гени в одну групу.

Про значення білка LEP див. в розділі 3.6.4 Кластер 4. Сигналінг через JAK-STAT шлях. Знижена експресія *MUC20* послаблює захисний бар'єр і сприяє проникненню імунних клітин із кровотока у тканину; захисні функції відновлюються наприкінці вагітності. Білок PLAT кодує тканиноспецифічну протеазу, яка перетворює пламзіноген у плазімн, який розщеплює фібрінові кров'яні згустки і запобігає тромбоутворенню. Якщо в першій половині вагітності низька активність запобігає кровотечі, то у другій половині запобігає тромботичним процесам. Експресія трьох генів суттєво зростає у другій половині вагітності, але залишається нижче вихідного рівня в першому триместрі.

#### **Результати даного розділу опубліковані в наступних роботах:**

Lykhenko, O., Frolova, A., & Obolenska, M. (2017) "Designing the database for microarray experiments metadata", *2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF)*. IEEE. Available at: doi:10.1109/YSF.2017.8126658.

Lykhenko, O., Frolova, A., & Obolenskaya, M. (2017). Creation of gene expression database on preeclampsia-affected human placenta. *Biopolym. Cell*, 33(6), 442–452. doi: 10.7124/bc.000967

Lykhenko, O., Frolova, A., & Obolenskaya, M. (2021). Consecutive integration of available microarray data for analysis of differential gene expression in human placenta. *Biotechnologia Acta* 14(1), 38-45. Doi: <https://doi.org/10.15407/biotech14.01.38>

Lykhenko, O., Frolova, A., & Obolenskaya, M. (2021). Changes in the human placental transcriptome during the physiological course of pregnancy. *Biopolym. Cell*, 37(1), 73–82. doi: 10.7124/bc.000A4D

## РОЗДІЛ 4

### АНАЛІЗ І УЗАГАЛЬНЕННЯ РЕЗУЛЬТАТІВ

Обговорення розділено на чотири блоки, які стосуються створення бази метаданих для зразків матково-плацентарного комплексу, порівняльного аналізу генної експресії в плаценті між I і II і між II і III триместрами вагітності, особливостям «співробітництва» генів у плаценті та асоційованим з прееклампсією генам у складі диференційно експресованих генів.

#### **4.1. Створення бази метаданих для зразків матково-плацентарного комплексу в умовах нормального і патологічного перебігу вагітності**

Створення спеціалізованої бази метаданих IGEA, значно збагаченої за рахунок інформації, взятої з першоджерел, листування з авторами та з самих даних експресії (стать плоду), є зручним джерелом для різних варіантів порівняльного аналізу генної експресії за різними метаданими. На противагу GEO і ArrayExpress, вона надає можливість напряму шукати в датасетах зразки матково-плацентарного комплексу з потрібними характеристиками без необхідності завантажувати, вичитувати і стандартизувати метадані в кожному датасеті. База IGEA слугувала основою для проведення теперішнього дослідження і може бути використана для інших варіантів вторинного аналізу генної експресії в межах матково-плацентарного комплексу. Періодичне оновлення такої бази забезпечує її довгу життєздатність. На даний час база обмежена результатами, отриманими мікроарей-технологією і патологією, пов'язаною з прееклампсією. Збагачення бази метаданими, отриманими в результаті секвенування, і іншими варіантами

патології, пов'язаними з матково-плацентарним комплексом, є прийдешньою задачею.

#### **4.2. Порівняльний аналіз генної експресії в плаценті між I і II і між II і III триместрами фізіологічної вагітності**

Численні дослідження генної експресії використовують фізіологічну вагітність в якості контролю в певних часових точках дослідження і зрідка звертаються до детального дослідження всього перебігу фізіологічної вагітності, хоча такі системні дані є ґрунтовною основою для будь-яких порівняльних досліджень.

В роботі використаний інтегративний аналіз наявних відкритих даних з генної експресії, що надало можливість узагальнити напрацювання попередніх років і отримати кількісну характеристику експресії білок-кодуєчих генів у їх природному співвідношенні.

Генна експресія в плаценті впродовж вагітності відображає класичний процес розвитку органу з диференціюванням клітин, морфогенезом тобто утворенням тривимірної структури органу і ростом. До основних категорій розвитку плаценти доєднується імунна відповідь, до значної міри пов'язана з необхідністю толерувати алогенний плід і запобігати його відторгненню і активізувати захисні функції на випадок атаки з боку зовнішніх шкідливих чинників і внутрішніх, які супроводжують процеси диференціювання і активного росту. **Мета роботи полягала в тому, щоб наповнити ці загальновідомі категорії розвитку генами, притаманними плаценті на часовій шкалі її розвитку.**

*Експресія генів, задіяних у морфогенезі.* Перша половина вагітності характеризується активацією морфогенетичних процесів, формуванням структури

і форми органу, диференціюванням клітин, васкуло- й ангиогенезом, зростанням ворсинок хоріону і проникненням позаворсинчастого трофобласту у децидуальну оболонку матері. Різні аспекти цих процесів ілюструються і деталізуються збагаченими біологічними шляхами і відповідними генами, які входять до класичних каскадів морфогенів, Wnt, Hippo, Notch і ретиноевої кислоти, гомеотичних транскрипційних факторів і факторів комітування. В першій половині експресія відповідних генів активована, у другій вона йде на спад (рис. 3.10, рис. 3.21).

*Експресія генів комітування*, які зумовлюють диференціювання клітин у певному напрямку і визначають подальшу специфікацію клітин. Система Gene Ontology (GO) визначає збагачені біологічні процеси як ті, які відповідають за створення певних органів і структур організму як, наприклад, підшлункової залози, кінцівок, легенів, тощо. Експресія цих генів (*GATA6*, *HAND1*, *ISL1*, *PITX2*) суттєво знижується у другій половині вагітності. Ті ж самі гени присутні у матриці експресії в періоді 1\_2, але вони не потрапили до диференційно експресованих з  $|\text{LogFC}| > 1$ , окрім гена *ISL1* (1,35). Значна частина цих генів — це транскрипційні фактори, і було не зрозуміло яким чином реалізується їх специфічна дія в плаценті і в процесі ембріогенезу. Перше пояснення з'явилося у публікації 2025 р., в якій доводиться, що експресія гена *HAND1* притаманна стовбуровим клітинам, а напрямок їх розвитку (комітування або спеціалізація) залежить від інтенсивності експресії гена і кількості білка, який клітини продукують. При низькій інтенсивності експресії гена і відповідно низькій концентрації білка стовбурові клітини диференціюються в клітини-попередники кардіоміоцитів і епікардіальних клітин, а при високій концентрації білка — в клітини екстраембріональної мезодерми [114]. Ген *PITX2*, як і *HAND1* ген, також є маркером позаембріональної мезодерми, належить до бікоїд-подібних генів, які

обумовлюють вісі тіла. Він бере участь у формуванні численних органів і найбільш вивчена його участь у формуванні структур серця (шлуночків і т.п.), лівостороннього розташування серця і навіть ритму серцебиття [115]. В плаценті він бере активну участь у процесі імплантації [116]. Стосовно двох генів, **ISL1** і **GATA6**, такого стибу питань не виникає через те, що на сьогодні доведено, що **ISL1** причетний до морфогенезу саме плаценти, її судинної системи і амніону [117], а **GATA6** знижує плюрипотентність стовбурових клітин і стимулює експресію екстраембріональних генів ентодерми [118].

Окрім зниженої експресії перелічених генів комітування у періоді 2\_3 спостерігається підвищення експресії двох генів із надродина TGF-beta, пептидних гормонів **INHBA** і **INHA**, які інгібуючи активіни блокують диференціювання і синтез колагену.

*Експресія генів, які забезпечують зростання органу і імунну толерантність.* Експресія генів, задіяних у проліферативних процесах і в імунній відповіді (прозапальні цитокіни) змінюється за такою самою схемою, як і експресія морфогенів, тобто зростає в першій половині вагітності й спадає у другій (рис. 3.7, 3.15). Припускаємо, що прозапальна реакція активується на пошкодження в процесі імплантації та на відпрацьовані клітини на тлі створення нових. Прозапальні реакції можуть також захищати тканину від можливих екзогенних чинників у критичні і найбільш незахищені періоди розвитку. Наприклад, гени **PAEP** (progestagen associated endometrial protein, -4,43) і **DEFB1** (defensin beta 1, -3,23) необхідні на стадії імплантації, запобігають відторгненню бластоцисти і захищають клітини від дії внутрішніх і можливих зовнішніх чинників; в подальшому їх активність не потрібна і експресія генів спадає.

*Варіанти ко-експресії генів впродовж вагітності.* Перетин диференційно експресованих генів в першій і другій половині вагітності з наступною

нормалізацією даних дозволив виокремити чотири групи ко-експресованих генів. У двох групах експресія генів або невпинно зростає або невпинно спадає впродовж вагітності і у двох групах експресія генів змінює направленість на протилежну при переході від першої половини до другої половини вагітності. Таким чином варіабельність ко-експресії генів мінімізується до чотирьох типів і при цьому, передбачаємо, суттєво економляться енергетичні ресурси клітини/тканини на регуляцію експресії. Численні фактори можуть бути задіяні у ко-експресії наборів генів у просторі і часі: утворення оперонів, топологічно пов'язаних доменів і транскрипційних конденсатів, використання спільних енансерів через утворення петель ДНК, використання двонаправлених промоторів, кластеризація генів та формування топологічно пов'язаних доменів [119]. Відомо, що біля 87% транскриптів людського геному є коекспресованими. Цей цікавий аспект регуляції генної активності потребує подальших досліджень.

Переломний момент в експресії генів при переході від другого до третього триместру наведено на багатьох прикладах у роботі і особливо наочно на прикладі двох типів ко-експресованих генів (рис. 3.26, 3.28). Цікаво, що цей перелом співпадає в часі з першою появою клінічних ознак ускладнення вагітності преєклампсією, і питання полягає в тому, чому саме в цей момент вперше з'являються клінічні ознаки, не раніше і не пізніше. Ми припускаємо наступний розвиток подій. На початку вагітності зростання маси плаценти випереджає зростання маси плоду, і плацента здатна забезпечити потреби іще дуже маленького ембріона навіть в разі її некритичної дисфункції, яка в подальшому дає себе взнаки. На початку другого триместру плід починає набирати масу значно інтенсивніше, ніж плацента, і цей розрив весь час збільшується до кінця вагітності (рис. 4.1). В умовах недостатності плаценти це дається взнаки у вигляді клінічних проявів. Переломний момент в експресії генів зафіксований всередині гестації в

базальній пластини плаценти людини [120] і в плацентах мишей при незмінних клітинному складі і морфології плаценти (Knox et al., 2008).

Щодо молекулярних механізмів цього переходу, то заслуговують на увагу гени, які різко змінюють направленість експресії і механізми регуляції їх експресії. Ключову роль в цьому переході можуть відігравати: зміна в наборі транскрипційних факторів (HIF1 $\alpha$ , GCM1, STATs, та PPAR $\gamma$ ); епігенетична модифікація через метилування і дію мікро РНК, які впливають на стабільність мРНК; гормональна регуляція (hCG, progesterone, та estrogens), наприклад, рівень hCG високий на ранньому етапі гестації і спадає до середини, гормони регулюють сигнальні шляхи MAPK, PI3K/АКТ, та mTOR; зміна у клітинному наборі і диференціюванні клітин від проліферативного до високо диференційованого, і, нарешті, зміни в імунному статусі від прозапального до імунотолерантного і в наборі цитокінів.

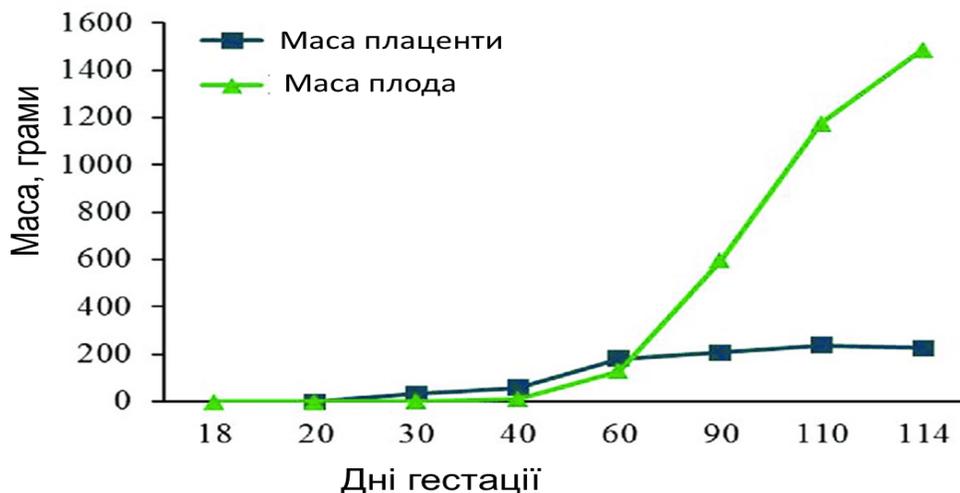


Рис. 4.1. Інтенсивність набору маси плодом (прямокутники) і плацентою (трикутники). Після середини гестації маса плоду збільшується швидше, ніж маса плаценти й ця «ножицеподібна» різниця збільшується до кінця вагітності. Такий розрив посилює чутливість системи до будь-якої дисфункції плаценти і є одним з пояснень, чому клінічні ознаки виявляються саме після 20 тижнів [128].

### 4.3. Особливості «співробітництва» генів в плаценті та між плацентою і плодом

Аналізуючи експресію численних генів з різними функціями, хочеться звернути увагу на декілька варіантів “співробітництва” генів для досягнення ефективного результату в межах певного біологічного процесу. Деякі гени співпрацюють за принципом взаємовиключення. Показовий приклад — активація прозапальних хемокінів з одночасним інгібуванням гомеостатичних хемокінів (рис. 3.8). Інші гени синергічно “йдуть протилежними шляхами” для досягнення спільної мети. Наприклад, у порівнянні 1\_2 (кластер 2) збільшення експресії позитивних чинників (активаторів) морфогенезу супроводжується одночасним зменшенням експресії негативних чинників (інгібіторів), що забезпечує найбільшу спрямовану ефективність морфогенетичного процесу. У порівняння 2\_3 зменшення експресії генів комітування супроводжується активацією інгібінів, які інгібують активіни і через них блокують диференціювання клітин (кластер 3).

Співробітництво інших генів має більш складний характер. Наприклад, зниження у другій половині вагітності активності гена *AOC1* (рис. 3.26), єдиного ензима, який розщеплює гістамін, створює умови для зниження тону судин і зменшення артеріального тиску, бо менше гістаміну розщеплюватиметься. Одночасно знижується експресія гена *REN* (рис. 3.27), який кодує високоспецифічну ендопептидазу, єдиною відомою функцією якої є перетворення ангіотензиногену в ангіотензин, який у свою чергу ініціює каскад реакцій, що призводять до підвищення артеріального тиску та збільшення затримки натрію нирками. Тобто синергічне зниження активності генів *AOC1* і *REN* спрямоване на зниження артеріального тиску. З іншого боку, на цьому етапі знижується експресія гена *PRG2* (рис. 3.26), що сприятиме вивільненню ангіотензиногена із комплексу з *PRG2*, і вірогідному перетворенню звільненого ангіотензиногену в ангіотензині

підвищенню артеріального тиску. До дії цих трьох білків приєднуються нейротензин (*NTS*), тридекапептид (складається з 13 амінокислот), вазоактивний пептид адреномедулін (*ADM*) і ендотеліальна NO-синтаза (*NOS3*, 1,78). Обидва мають широкий вплив на серцево-судинну систему, включаючи здатність знижувати/підвищувати артеріальний тиск (*NTS*) [125, 126] і знижувати артеріальний тиск, розслаблюючи гладкі м'язи судин (*ADM* і *NO*). Тобто порушення балансу щонайменше між шістьма цими гравцями може призводити до першого клінічного прояву преєклампсії — підвищенню артеріального тиску.

Інший характер співробітництва встановлюється між плацентою і плодом. Впродовж усього періоду вагітності в плаценті неухильно зменшується експресія генів альфа-фетопротейну (*AFP*), який виконує функції ембріонального альбуміна, і гена *HBE1*, який кодує епсилон ( $\epsilon$ ) субодиницю фетального гемоглобіну (рис. 3.28). Функції обох перебирає на себе печінка плоду.

#### **4.4. Диференційно експресовані гени, асоційовані з преєклампсією**

Як було згадано раніше, ряд ДЕГів у цьому дослідженні відомі також як маркери преєклампсії, що не дивує, бо ними можуть бути будь-які гени з порушеною регуляцією їх експресії, білкові ізоформи або — як у разі з відомим маркером преєклампсії sFLT — екстраклітинною розчинною ділянкою рецептора VEGF1. Ми порівняли диференційно експресовані гени між першим і другим (253 ДЕГи) та другим і третім (489 ДЕГів) триместрами вагітності з генами з бази даних порівняльної токсикогеноміки (Comparative Toxicogenomics Database, CTD) [123], а саме з переліком генів у розділі “Асоціація генів із захворюваннями”. Транскриптом людини налічує 376 генів, асоційованих з терміном “преєклампсія” (MeSH term Pre-Eclampsia, D011225) [123]. Внаслідок порівняння ДЕГів з цим

переліком виявили 23 ДЕГи у порівнянні 1\_2 і 38 ДЕГів у порівнянні 2\_3, які асоційовані з прееклампсією (Додаток Д). Серед них гени, що впливають на тонус судин (REN), згортання крові (SERPINB2), запалення (IL6) та хемотаксис.

Ми побудували мережу білок-білкових взаємодій і асоціацій для білків, які кодуються генами, асоційованими з прееклампсією, у програмі String.db. Гени з періоду 1\_2 пов'язані з дією прозапального цитокіну ІЛ-17 (рис. 4.2), а гени з періоду 2\_3 з дією протизапального цитокіна ІЛ\_10 (рис. 4.3).

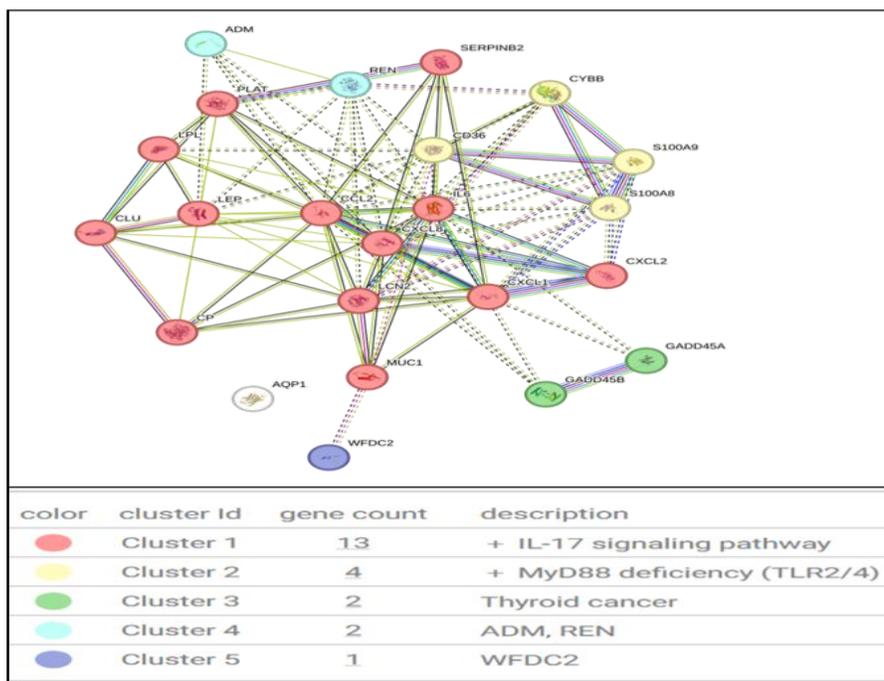


Рис. 4.2. Графи білок-білкових взаємодій та асоціацій відповідно до асоційованих з прееклампсією диференційно експресованих генів з періоду 1\_2

ІЛ-17 належить до групи цитокінів, відомих переважно як інтерлейкін-17А (ІЛ-17А), які відіграють вирішальну роль в імунній системі, зокрема в запаленні та захисті організму. Він синтезується різними імунними клітинами, включаючи клітини Th17 і синцитіотрофобласт, і бере участь як у природних імунних реакціях захисту організму від різних несприятливих чинників, так і в розвитку запальних

реакцій. Подвійний характер активності ІЛ-17 відповідає процесам, які відбуваються у першій половині вагітності. Цікаво, що ІЛ-17 може проходити через плаценту до плоду, але це стосується материнського ІЛ-17, який в певних умовах виробляється у значній кількості і, проходячи через плаценту, негативно впливає на розвиток центральної нервової системи плоду [129]. ІЛ-10, потужний протизапальний цитокін, відіграє вирішальну роль у підтримці імунної толерантності у ділянці материнсько-плацентарного зіткнення. ІЛ-10 виробляється трофобластами (впливаючи на материнські децидуальні клітини) та материнськими імунними клітинами в децидуальній оболонці (впливаючи на трофобласт або інші клітини фетального походження) [130].

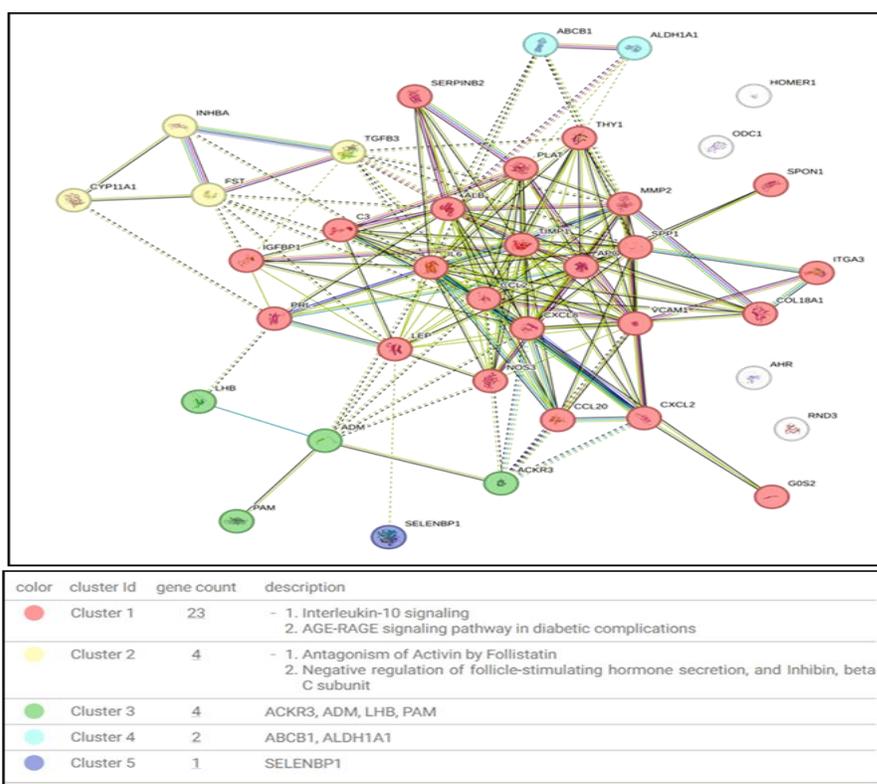


Рис. 4.3. Графи білок-білкових взаємодій та асоціацій відповідно до асоційованих з преєклампсією диференційно експресованих генів з періоду 2\_3

За даними НРА значна кількість відповідних білків секретується: в періоді

1\_2 — хемокіни **CXCL1**, **CXCL2**, **CXCL8 (IL-8)**, цитокин **IL6**, **LCN2** (Lipocalin-2) і **PLAT** (тканинний активатор плазміногену); в періоді 2\_3 — 26 білків з 53х є секреторними білками (див. Додаток Д). Через молекули, комплекси і структури, які секретуються, плацента взаємодіє з організмом матері, а в разі прееклампсії призводить до генералізованого ураження організму матері.

Проведене дослідження відповідає принципам системної біології. Воно є міждисциплінарним, бо поєднує дані молекулярної біології (з баз даних GEO, ArrayExpress) з обчислювальними методами комп'ютерних наук та математичними статистичними методами. У використаних даних вимірюється експресія не окремих генів, а визначається профіль генної експресії в цілому. Кількісні дані про зміну експресії генів між триместрами поєднуються із загальними даними про залучення цих генів до біологічних процесів та до клітинних компонентів у людському організмі (база Gene Ontology), мережових міжбілкових взаємодій (база STRING), секретому (база Human Protein Atlas), а також асоціацій із патологічними станами, зокрема, прееклампсією (база CTD). Тому інтегративність аналізу проявляється не лише в об'єднанні однорідних датасетів, а й у поєднанні різнорідних баз біологічних даних, що відображають регуляцію, взаємодію та локалізацію генів. Такий підхід дозволяє оцінити функціонування біологічної системи на молекулярному, клітинному та на рівні організму і втілює принципи системної біології, спрямованої на побудову цілісного уявлення про найскладнішу систему людського організму: мати – плацента – плід.

## ВИСНОВКИ

Створено унікальну спеціалізовану базу метаданих для вторинного аналізу транскриптому у зразках матково-плацентарного комплексу та крові матері й дитини. Проведено системний інтегративний аналіз генної експресії в плаценті, який виявив загальні зміни між першим і другим та другим і третім триместрами нормальної вагітності, а також закономірності ко-експресії генів впродовж вагітності та взаємодії генів у межах одного біологічного процесу.

1. Створено спеціалізовану базу стандартизованих метаданих (IGEA) з 48 датасетів, яка налічує 1466 зразків для зразків матково-плацентарного комплексу та крові матері й дитини, яка уможливорює простий та швидкий відбір даних за обраними користувачем характеристиками зразків.
2. На підставі даних експресії генів, що були відібрані з обраних з бази IGEA датасетів восьми досліджень цілісної тканини плаценти першого, другого та третього триместрів, було утворено новий інтегрований датасет. За допомогою аналізу диференційної експресії в інтегрованому датасеті отримано оригінальні дані, що кількісно характеризують зміни в експресії 253 генів між першим і другим та 489 генів між другим і третім триместрами вагітності. Встановлено, що, перший часовий інтервал характеризується активними імунними і захисними реакціями, диференціюванням клітин і ростом органу, прозапальними процесами і підтриманням взаємної толерантності у ділянці зіткненням плодової частини плаценти з децидуальною тканиною матері. У свою чергу, другому інтервалу притаманне суттєве зменшення прозапальних процесів, на заміну яким приходять протизапальні процеси, зменшення проліферативної активності й

активності морфогенів, генів комітування та активація метаболічних процесів, характерних для високо-диференційованої тканини.

3. Завдяки порівняльному аналізу диференційно експресованих генів між першим і другим та другим і третім триместрами встановлено 54 гени, які змінюють значення експресії на обох інтервалах, і вперше виокремлено чотири типи ко-експресії генів впродовж вагітності: експресія генів неухильно зростає впродовж усього періоду вагітності (18 генів), неухильно знижується від першого до другого і від другого до третього триместру (12 генів) і гени, траєкторія експресії яких змінюється всередині вагітності від збільшення у першій половині до зменшення у другій половині вагітності (21 ген) і навпаки (3 гена).
4. Вперше виокремлено типи взаємодії диференційно експресованих генів, які функціонують в межах одного біологічного процесу: взаємне виключення (активація прозапальних хемокінів (11 генів) супроводжується зниженням експресії гомеостатичних хемокінів (4 гени) у першій половині вагітності) і синергія (активація активаторів морфогенезу (12 генів) з одночасним інгібуванням їх інгібіторів (9 генів) у першій половині вагітності або активація субодиниць інгібіну (2 гени) при інгібуванні активіну).
5. Вперше впродовж розвитку плаценти встановлено зміну експресії генів комітування, які є активними в першій половині (6 генів) та знижують експресію у другій половині вагітності (5 генів).
6. Вперше серед диференційно експресованих генів між першим і другим та другим і третім триместрами нормальної вагітності визначено гени, які асоційовані з преєклампсією, і які з них кодують білки, що секретуються в кров.

7. Аналіз диференційної експресії генів в отриманому в даному дослідженні інтегрованого датасету продемонстрував зниження експресії плеiotропного маркерного гена лептину (*LEP*) у першій половині вагітності, яке вважалося суперечливим.

**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ**

- [1] Brosens, I., Pijnenborg, R., Vercruyssen, L. and Romero, R. (2011) "The "Great Obstetrical Syndromes" are associated with disorders of deep placentation", *American Journal of Obstetrics and Gynecology*, 204(3), pp. 193–201. Available at: doi:10.1016/j.ajog.2010.08.009.
- [2] Knöfler, M., Haider, S., Saleh, L., Pollheimer, J., Gamage, T.K.J.B. and James, J. (2019) "Human placenta and trophoblast development: key molecular mechanisms and model systems", *Cellular and Molecular Life Sciences*, 76(18), pp. 3479–3496. Available at: doi:10.1007/s00018-019-03104-6.
- [3] Ortega, M.A. *et al.* (2022) The pivotal role of the placenta in normal and pathological pregnancies: a focus on preeclampsia, fetal growth restriction, and maternal chronic venous disease. *Cells*, 11(3), p. 568. Available at: doi:10.3390/cells11030568.
- [4] Napso, T., Yong, H.E.J., Lopez-Tello, J. & Sferruzzi-Perri, A.N. (2018) The role of placental hormones in mediating maternal adaptations to support pregnancy and lactation. *Frontiers in Physiology*, 9, p. 1091. Available at: doi:10.3389/fphys.2018.01091.
- [5] Murphy, V.E., Smith, R., Giles, W.B. & Clifton, V.L. (2006) Endocrine regulation of human fetal growth: the role of the mother, placenta, and fetus. *Endocrine Reviews*, 27(2), pp. 141–169. Available at: doi:10.1210/er.2005-0011.
- [6] Gabory, A., Roseboom, T.J., Moore, T., Moore, L.G. & Junien, C. (2013) "Placental contribution to the origins of sexual dimorphism in health and diseases: sex chromosomes and epigenetics", *Biology of Sex Differences*, 4(1), p. 5. Available at: doi:10.1186/2042-6410-4-5.

- [7] Faller, A. & Schünke, M. (eds.) (2004) *The Human Body: An Introduction to Structure and Function*, 1st edn. Stuttgart: Thieme. Available at: doi:10.1055/b-005-148851.
- [8] Tal, R., Shaikh, S., Pallavi, P., Tal, A., López-Giráldez, F., Lyu, F. *et al.* (2019) Adult bone marrow progenitors become decidual cells and contribute to embryo implantation and pregnancy. *PLoS Biology*, 17(9), e3000421. Available at: doi:10.1371/journal.pbio.3000421.
- [9] Creighton, C. (1878) "The formation of the placenta in the guinea-pig", *Journal of Anatomy and Physiology*, 12(Pt 4), pp. 534–590.
- [10] Jaremek, A., Jeyarajah, M.J., Jaju Bhattad, G. & Renaud, S.J. (2021) "Omics approaches to study formation and function of human placental syncytiotrophoblast", *Frontiers in Cell and Developmental Biology*, 9, p. 674162. Available at: doi:10.3389/fcell.2021.674162.
- [11] Riddell, M.R., Winkler-Lowen, B., Chakrabarti, S., Dunk, C., Davidge, S.T. & Guilbert, L.J. (2012) The characterization of fibrocyte-like cells: a novel fibroblastic cell of the placenta. *Placenta*, 33, pp. 143–150.
- [12] James, J.L., Lissaman, A., Nursalim, Y.N.S. & Chamley, L.W. (2022) "Modelling human placental villous development: designing cultures that reflect anatomy", *Cellular and Molecular Life Sciences*, 79(7), pp. 1–20. Available at: doi:10.1007/s00018-022-04407-x.
- [13] Hempstock, J., Cindrova-Davies, T., Jauniaux, E. *et al.* (2004) "Endometrial glands as a source of nutrients, growth factors and cytokines during the first trimester of human pregnancy: a morphological and immunohistochemical study", *Reproductive Biology and Endocrinology*, 2, p. 58. Available at: doi:10.1186/1477-7827-2-58.

- [14] Carter, A.M. (2009) "Evolution of factors affecting placental oxygen transfer", *Placenta*, 30, pp. 19–25. Available at: doi:10.1016/j.placenta.2008.11.006.
- [15] Clifton, V.L. (2010) "Review: sex and the human placenta: mediating differential strategies of fetal growth and survival", *Placenta*. Available at: <https://pubmed.ncbi.nlm.nih.gov/20004469> (Accessed: 3 July 2025).
- [16] Moritz, K.M., Cuffe, J.S.M., Wilson, L.B., Dickinson, H., Wlodek, M.E., Simmons, D.G. & Denton, K.M. (2010) Review: sex-specific programming: a critical role for the renal renin–angiotensin system. *Placenta*. Available at: <https://pubmed.ncbi.nlm.nih.gov/20116093> (Accessed: 3 July 2025).
- [17] Di Renzo, G.C., Rosati, A., Sarti, R.D., Cruciani, L. & Cutuli, A.M. (2007) "Does fetal sex affect pregnancy outcome?", *Gender Medicine*, 4(1), pp. 19–30. Available at: doi:10.1016/S1550-8579(07)80004-0.
- [18] Peacock, J.L., Marston, L., Marlow, N., Calvert, S.A. & Greenough, A. (2012) Neonatal and infant outcome in boys and girls born very prematurely. *Pediatric Research*, 71(3), pp. 305–310. Available at: doi:10.1038/pr.2011.50.
- [19] Wallace, J.M., Horgan, G.W. & Bhattacharya, S. (2012) Placental weight and efficiency in relation to maternal body mass index and the risk of pregnancy complications in women delivering singleton babies. *Placenta*. Available at: <https://pubmed.ncbi.nlm.nih.gov/22695104> (Accessed: 3 July 2025).
- [20] Byrne, J. & Warburton, D. (1987) "Male excess among anatomically normal fetuses in spontaneous abortions", *American Journal of Medical Genetics*, 26(3), pp. 605–611. Available at: doi:10.1002/ajmg.1320260315.
- [21] Mao, J., Zhang, X., Sieli, P.T., Falduto, M.T., Torres, K.E. & Rosenfeld, C.S. (2010) "Contrasting effects of different maternal diets on sexually dimorphic gene expression in the murine placenta", *Proceedings of the National Academy*

of Sciences of the United States of America. Available at: <https://pubmed.ncbi.nlm.nih.gov/20212133> (Accessed: 3 July 2025).

- [22] Skinner, M.K. (2008) What is an epigenetic transgenerational phenotype? F3 or F2. *Reproductive Toxicology*, 25(1), p. 2. Available at: doi:10.1016/j.reprotox.2007.09.001.
- [23] Peleg, D., Kennedy, C.M. & Hunter, S.K. (1998) Intrauterine growth restriction: identification and management. *American Family Physician*, 58(2), p. 453. Available at: <https://www.aafp.org/afp/1998/0801/p453.html> (Accessed: 3 July 2025).
- [24] Baergen, R.N. (2012) "Placental Pathology in Pregnancy Complications", *Thrombosis Research*, 130(Suppl 1), pp. S51–S54. Available at: doi:10.1016/S0049-3848(12)70015-2.
- [25] Hill, M.A. (2021) "Embryology fetal development", *Embryology: Fetal Development*. Available at: [https://embryology.med.unsw.edu.au/embryology/index.php/Fetal\\_Development](https://embryology.med.unsw.edu.au/embryology/index.php/Fetal_Development) (Accessed: 25 May 2021).
- [26] Heindel, J.J. & Vandenberg, L.N. (2015) "Developmental origins of health and disease: a paradigm for understanding disease cause and prevention", *Current Opinion in Pediatrics*, 27(2), pp. 248–253. Available at: doi:10.1097/MOP.0000000000000191.
- [27] Benny, P.A., Alakwaa, F.M., Schlueter, R.J., Lassiter, C.B. & Garmire, L.X. (2020) "A review of omics approaches to study preeclampsia", *Placenta*, 92, pp. 17–27. Available at: doi:10.1016/j.placenta.2020.01.008.
- [28] Cirkovic, A. *et al.* (2020) "Systematic review supports the role of DNA methylation in the pathophysiology of preeclampsia: a call for analytical and

- methodological standardization", *Biology of Sex Differences*, 11, p. 36. Available at: doi:10.1186/s13293-020-00313-8.
- [29] Yong, H.E.J. & Chan, S.-Y. (2020) Current approaches and developments in transcript profiling of the human placenta. *Human Reproduction Update*, 26(6), pp. 799–840. Available at: doi:10.1093/humupd/dmaa028.
- [30] Uitert, M., Moerland, P.D., Enquobahrie, D.A. *et al.* (2015) Meta-analysis of placental transcriptome data identifies a novel molecular pathway related to preeclampsia. *PLoS One*, 10, e0132468. doi: 10.1371/journal.pone.0132468.
- [31] Souza, R.T., Mayrink, J., Leite, D.F., Costa, M.L., Calderon, I.M. & Rocha Filho, E.A. *et al.* (2019) Metabolomics applied to maternal and perinatal health: a review of new frontiers with translation potential. *Clinics*, 74, e894. Available at: doi:10.6061/clinics/2019/e894.
- [32] Kleinrouweler, C.E., van Uitert, M., Moerland, P.D., Ris-Stalpers, C., van der Post, J.A.M. & Afink, G.B. (2013) "Differentially expressed genes in the pre-eclamptic placenta: a systematic review and meta-analysis", *PLoS One*, 8(7), e68991. Available at: doi:10.1371/journal.pone.0068991.
- [33] Brew, O., Sullivan, M.H.F. & Woodman, A. (2016) "Comparison of normal and pre-eclamptic placental gene expression: a systematic review with meta-analysis", *PLoS One*, 11(8), e0161504. Available at: doi:10.1371/journal.pone.0161504.
- [34] Vaiman, D. & Miralles, F. (2016) An integrative analysis of preeclampsia based on the construction of an extended composite network featuring protein–protein physical interactions and transcriptional relationships. *PLoS One*, 11(11), e0165849. Available at: doi:10.1371/journal.pone.0165849.

- [35] Leavey, K., Bainbridge, S.A. & Cox, B.J. (2015) "Large scale aggregate microarray analysis reveals three distinct molecular subclasses of human preeclampsia", *PLoS One*, 10(2), e0116508. doi: 10.1371/journal.pone.0116508.
- [36] Liu, H., Bebu, I. & Li, X. (2010) "Microarray probes and probe sets", *Frontiers in Bioscience (Elite edition)*, 2, pp. 325–338. Available at: doi:10.2741/e93.
- [37] Miller, J.A., Cai, C., Langfelder, P., Geschwind, D.H., Kurian, S.M., Salomon, D.R. & Horvath, S. (2011) "Strategies for aggregating gene expression data: the collapseRows R function", *BMC Bioinformatics*, 12(1), pp. 1–13. Available at: doi:10.1186/1471-2105-12-322.
- [38] Frolova, A., Bondarenko, V. & Obolenska, M. (2016) "Comparing alternative pipelines for cross-platform microarray gene expression data integration with RNA-seq data in breast cancer", *bioRxiv*. Available at: doi:10.1101/059600.
- [39] Turnbull, A.K., Kitchen, R.R., Larionov, A.A., Renshaw, L., Dixon, J. & Sims, A.H. (2012) Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. *BMC Medical Genomics*, 5(1), p. 35.
- [40] Tseng, G.C., Ghosh, D. & Feingold, E. (2012) Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Research*, 40(9), pp. 3785–3799.
- [41] Gautier, L., Cope, L., Bolstad, B.M. and Irizarry, R.A. (2004) "affy—analysis of Affymetrix GeneChip data at the probe level", *Bioinformatics*, 20(3), pp. 307–315. Available at: doi:10.1093/bioinformatics/btg405.
- [42] Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003) "Summaries of Affymetrix GeneChip probe level data", *Nucleic Acids Research*, 31(4), p. e15.

- [43] Du, P., Kibbe, W.A. and Lin, S.M. (2008) "lumi: a pipeline for processing Illumina microarray", *Bioinformatics*, 24(13), pp. 1547–1548. Available at: doi:10.1093/bioinformatics/btn224 [bioconductor.org](http://bioconductor.org)
- [44] Venables, W.N. and Ripley, B.D. (2002) *Modern Applied Statistics with S*. New York: Springer-Verlag, p. 562.
- [45] Van Der Maaten, L.J.P. and Hinton, G.E. (2008) "Visualizing High-Dimensional Data Using t-SNE", *Journal of Machine Learning Research*, 9(Nov), pp. 2579–2605.
- [46] Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015) "limma powers differential expression analyses for RNA-sequencing and microarray studies", *Nucleic Acids Research*, 43(7), p. e47. Available at: doi:10.1093/nar/gkv007.
- [47] Buckberry, S. *et al.* (2014) "massiR: a method for predicting the sex of samples in gene expression microarray datasets", *Bioinformatics*, 30(14), pp. 2084–2085. Available at: doi:10.1093/bioinformatics/btu161.
- [48] Benjamini, Y. and Hochberg, Y. (1995) "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing", *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), pp. 289–300. Available at: <http://www.jstor.org/stable/2346101> (Accessed: 3 July 2025).
- [49] Smyth, G.K. (2005) "limma: Linear models for microarray data", in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, pp. 397–420. doi:10.1007/0-387-29362-0\_23.
- [50] Szklarczyk, D. *et al.* (2019) "STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets", *Nucleic Acids Research*, 47(Database), pp. D607–D613. Available at: doi:10.1093/nar/gky1131.

- [51] Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R. *et al.* (2023) "The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest", *Nucleic Acids Research*, 51(D1), pp. D638–D646. Available at: doi:10.1093/nar/gkac1000.
- [52] Franceschini, A *et al.* (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. In: *Nucleic Acids Res.* 2013 Jan;41(Database issue):D808-15. doi: 10.1093/nar/gks1094.
- [53] Podani, J. (2021) "The wonder of the Jaccard coefficient: from alpine floras to bipartite networks", *Flora Mediterranea*, 31(Special Issue), pp. 105–123. Available at: doi:10.7320/FlMedit31SI.105.
- [54] Maleki, F. and Kusalik, A.J. (2019) "Gene set overlap: an impediment to achieving high specificity in over-representation analysis", *bioRxiv*. Preprint. Available at: doi:10.1101/319145.
- [55] Leavey, K., Bainbridge, S. A., & Cox, B. J. (2015). Large Scale Aggregate Microarray Analysis Reveals Three Distinct Molecular Subclasses of Human Preeclampsia. *PLoS One*, 10(2), e0116508. Available at: doi:10.1371/journal.pone.0116508.
- [56] Zhu, L. *et al.* (2018) Bacterial communities in the womb during healthy pregnancy. *Frontiers in Microbiology*, 9, p. 2163. Available at: doi:10.3389/fmicb.2018.02163.
- [57] Hughes, C.E., Nibbs, R.J.B. (2018). A guide to chemokines and their receptors. *FEBS J* 285, pp. 2944-2971. Available at: doi:10.1111/febs.14466.
- [58] Kim, W., Khan, S.K., Gvozdenovic-Jeremic, J., Kim, Y., Dahlman, J., Kim, H., Park, O., Ishitani, T., Jho, E.H., Gao, B. and Yang, Y. (2017) "Hippo signaling interactions with Wnt/ $\beta$ -catenin and Notch signaling repress liver

- tumorigenesis", *Journal of Clinical Investigation*, 127(1), pp. 137-152. Available at: doi:10.1172/JCI88486.
- [59] Murthi, P., Abumaree, M. and Kalionis, B. (2014) "Analysis of homeobox gene action may reveal novel angiogenic pathways in normal placental vasculature and in clinical pregnancy disorders associated with abnormal placental angiogenesis", *Frontiers in Pharmacology*, 5, article 133. Available at: doi:10.3389/fphar.2014.00133.
- [60] Katagiri, T. and Watabe, T. (2016) "Bone Morphogenetic Proteins", *Cold Spring Harbor Perspectives in Biology*, 8, a021899. Available at: doi:10.1101/cshperspect.a021899.
- [61] Wang, H., Fan, N., Cui, X., Xie, R., Tang, Y., Thomas, A. M., Li, S., Zhang, J. V., Liu, S. & Qin, H. (2024) BMP5 promotes trophoblast functions upon N-glycosylation via the BMP5-SMAD1/5 signaling pathway in preeclampsia. *Placenta*, 158, pp. 240–252. Available at: doi:10.1016/j.placenta.2024.11.002.
- [62] Somekawa, S., Imagawa, K., Hayashi, H., Sakabe, M., Ioka, T., Sato, G. E., Inada, K., Iwamoto, T., Mori, T., Uemura, S., Nakagawa, O. & Saito, Y. (2012) Tmem100, an ALK1 receptor signaling-dependent gene essential for arterial endothelium differentiation and vascular morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 109(30), pp. 12064–12069. Available at: doi:10.1073/pnas.1207210109.
- [63] Ikeno, S., Nakano, N., Sano, K., Minowa, T., Sato, W., Akatsu, R., Sakata, N., Hanagata, N., Fujii, M., Itoh, F. & Itoh, S. (2019) PDZK1-interacting protein 1 (PDZK1IP1) traps Smad4 protein and suppresses transforming growth factor- $\beta$  (TGF- $\beta$ ) signaling. *Journal of Biological Chemistry*, 294(13), pp. 4966–4980. Available at: doi:10.1074/jbc.RA118.004153.

- [64] Kin Ting Kam, R., Deng, Y., Chen, Y., et al. (2012) "Retinoic acid synthesis and functions in early embryonic development", *Cell Bioscience*, 2, article 11. Available at: doi:10.1186/2045-3701-2-11.
- [65] Zheng, Y., Wu, L., Hu, Z. *et al.* Role of the Forkhead box family protein *FOXF2* in the progression of solid tumor: systematic review. *J Cancer Res Clin Oncol* 151, 14 (2025). <https://doi.org/10.1007/s00432-024-06047-z>
- [66] Lu, Y., Zhon, L., Luo, X., Liu, C., Dan, W., Chu, X., Liu, B. (2022) "MiRNA301b-3p induces proliferation and inhibits apoptosis in AML cells by targeting *FOXF2* and regulating Wnt/beta-catenin axis", *Mol Cell Probes* 63:101805. Available at: doiorg/101016/jmcp20221018 0 5.
- [67] Gauster, M., Desoye, G., Tötsch, M. and Hiden, U. (2007) "The placental lipid metabolism and its role in gestational diabetes mellitus", *Molecular Aspects of Medicine*, 28(1), pp. 90-101. Available at: doi:10.1016/j.mam.2006.12.005.
- [68] Yu, X., Meng, X., Xu, M., Zhang, S., Zhang, W. and Wang, X. (2018) "Lipoprotein lipase in placental tissue and maternal plasma during pregnancy and its relation to fetal growth", *Reproductive Biology and Endocrinology*, 16, article 43. Available at: doi:10.1186/s12958-018-0368-1.
- [69] Yang, R., Roshani, D., Gao, B., Li, P., Shang, N. (2024) "Metallothionein: A Comprehensive Review of Its Classification, Structure, Biological Functions, and Applications". *Antioxidants (Basel)*. 13(7):825. Available at: doi: 10.3390/antiox13070825.
- [70] Uchida, H., Maruyama, T., Nagashima, T., Asada, H., Yoshimura, Y. and Oda, H. (2013) "Progesterone-induced blocking factor (PIBF) mediates implantation and maintenance of pregnancy", *Reproductive Sciences*, 20(9), pp. 1020-1027. Available at: doi:10.1177/1933719113477485.

- [71] Lee, C.-L., Chiu, P.C.N., Hautala, L., Salo, T. and Yeung, W.S.B. (2016) "Progesterone-induced blocking factor (PIBF) promotes invasion of human extravillous trophoblasts and is up-regulated by interleukin-33", *Scientific Reports*, 6, article 32376. Available at: doi:10.1038/srep32376.
- [72] Liao, X.H., Wang, N., Zhang, S., Song, X., Zhou, W., Zhang, Y., Li, J. and Wang, X. (2018) "Placental Growth Hormone: An Important Regulator in Gestational Insulin Resistance, Lipid Metabolism and Fetal Growth", *Frontiers in Endocrinology*, 9, article 705. Available at: doi:10.3389/fendo.2018.00705.
- [73] Handwerger, S. (1991) "Placental lactogen, human chorionic somatomammotropin", *Endocrine Reviews*, 12(4), pp. 329-336. Available at: doi:10.1210/edrv-12-4-329.
- [74] King, B.R., Nicholson, R.C. and Smith, R. (2001) "Placental corticotrophin-releasing hormone, local effects and fetomaternal endocrinology", *Stress*. Available at: <https://pubmed.ncbi.nlm.nih.gov/22432143> (Accessed: 12 May 2025).
- [75] Hong, X., Wang, X., Wang, T., Zhu, H., Shao, J., Jiang, L. and Zhao, Y. (2019b) "Transcriptomic changes in human placenta in preeclampsia", *Placenta*, 85, pp. 1-8. Available at: doi:10.1016/j.placenta.2019.07.011.
- [76] Li, Y., Ma, L., He, R., Teng, F., Qin, X., Liang, X., Wang, J. (2024) "Pregnancy Metabolic Adaptation and Changes in Placental Metabolism in Preeclampsia. *Geburtshilfe Frauenheilkd*". 84(11):1033-1042. Available at: doi: 10.1055/a-2403-4855.
- [77] Karahoda, R., Kallol, S., Groessl, M., Ontsouka, E., Anderle, P., Fluck, C., Staud, F., Albrecht, C. (2021) "Revisiting Steroidogenic Pathways in the Human Placenta and Primary Human Trophoblast Cells". *Int J Mol Sci.*; 22(4):1704. Available at doi: 10.3390/ijms22041704.

- [78] Goldman-Wohl, D.S., Ariel, I., Greenfield, C., Hanoch, J. & Yagel, S. (2000) "HLA-G expression in extravillous trophoblasts is an intrinsic property of cell differentiation: a lesson learned from ectopic pregnancies", *Molecular Human Reproduction*, [online] PMID: 10825371.
- [79] Reisch, V., Gorbokov, N., Luebke, A.M., Burandt, E., Menz, A., Kluth, M. & Krech, T. (2021) "Pattern of placental alkaline phosphatase (PLAP) expression in human tumors: a tissue microarray study on 12,381 tumors", *Journal of Pathology: Clinical Research*, 7(6), pp. 577–589. doi: 10.1002/cjp2.237
- [80] Sola, I.M. *et al.* (2023) WNT5A,  $\beta$ -catenin and SUFU expression patterns, and the significance of microRNA deregulation in placentas with intrauterine growth restriction. *Molecular Medicine Reports*, 27(2), p. 28. Available at: doi:10.3892/mmr.2022.12914.
- [81] Xue, C., Chu, Q., Shi, Q. *et al.* (2025) "Wnt signaling pathways in biology and disease: mechanisms and therapeutic advances". *Sig Transduct Target Ther* 10, 106. Available at: [doi.org/10.1038/s41392-025-02142-w](https://doi.org/10.1038/s41392-025-02142-w).
- [82] Liu, Y. & Ma, Y. (2017) Promoter methylation status of WNT2 in placenta from patients with preeclampsia. *Medical Science Monitor*, 23, pp. 5294–5301. Available at: doi:10.12659/msm.903602.
- [83] Moore, T. and Dveksler, G.S. (2014) "Pregnancy-specific glycoproteins: complex gene families regulating maternal–fetal interactions", *International Journal of Developmental Biology*, 58(2-3-4), pp. 273–280. Available at: doi:10.1387/ijdb.140080gm.
- [84] Zhao, HJ., Klausen, C., Li, Y. *et al.* Bone morphogenetic protein 2 promotes human trophoblast cell invasion by upregulating N-cadherin via non-canonical SMAD2/3 signaling. *Cell Death Dis* 9, 174 (2018). <https://doi.org/10.1038/s41419-017-0230-1>

- [85] [Deng, J., Li, Y.](#) BMP6 Mediates BMP2-Increased Human Trophoblast Invasion. *J Endocr Soc.* 2021 May 3;5(Suppl 1):A747. doi: [10.1210/jendso/bvab048.1519](https://doi.org/10.1210/jendso/bvab048.1519)
- [86] Makanji, Y., Zhu, J., Mishra, R. et al. (2014) "Inhibin at 90: from discovery to clinical application, a historical review", *Endocrine Reviews*, 35(5), pp. 747–794. Available at: doi:10.1210/er.2014-1003.
- [87] Jenkin, G., Ward, J., Loose, J., Schneider-Kolsky, M., Young, R., Canny, B., O'Connor, A., de Kretser, D.M. and Wallace, E.M. (2001) "Physiological and regulatory roles of activin A in late pregnancy", *Molecular and Cellular Endocrinology*, 180(1–2), pp. 131–138. Available at: doi:10.1016/S0303-7207(01)00504-4.
- [88] Madazli, R., Kuseyrioglu, B., Uzun, H., Uludag, S. and Ocak, V. (2005) "Prediction of preeclampsia with maternal mid-trimester placental growth factor, activin A, fibronectin and uterine artery Doppler velocimetry", *International Journal of Gynaecology and Obstetrics*, 89(3), pp. 251–257. Available at: doi:10.1016/j.ijgo.2005.02.008.
- [89] Sugulle, M., Dechend, R., Herse, F., Weedon-Fekjaer, M.S., Johnsen, G.M., Brosnihan, K.B., Anton, L., Luft, F.C., Wollert, K.C., Kempf, T. and Staff, A.C. (2009) "Circulating and placental growth-differentiation factor 15 in preeclampsia and in pregnancy complicated by diabetes mellitus", *Hypertension*, 54(1), pp. 106–112. Available at: doi:10.1161/HYPERTENSIONAHA.109.130583.
- [90] Hussein, M.Y., Al-Salih, R.M.H. & Al-Nasser, A.H. (2025) A study of gene expression of inhibin and activin subunits (INH<sub>A</sub>, INH<sub>B</sub>A) in the placenta of women with missed miscarriage. *Cuest. Fisioter.*, 54(3), pp. 3460–3477.
- [91] Mylonas, I. et al. (2006) Expression of inhibin/activin subunits alpha ( $\alpha$ ), beta A ( $\beta$ A) and beta B ( $\beta$ B) in placental tissue of normal and intrauterine growth

- restricted (IUGR) pregnancies. *Journal of Molecular Histology*, 37, pp. 43–52. Available at: doi:10.1007/s10735-006-9029-6.
- [92] Muttukrishna, S. (2004) Role of inhibin in normal and high-risk pregnancy. *Seminars in Reproductive Medicine*, 22(3), pp. 227–234. Available at: doi:10.1055/s-2004-831898.
- [93] Fejzo, M. *et al.* (2024) GDF15 linked to maternal risk of nausea and vomiting during pregnancy. *Nature*, 625, pp. 760–767. Available at: doi:10.1038/s41586-023-06921-9.
- [94] HER Foundation (2024) "Genetics and GDF15 hormone FAQ", *HER Foundation* [online]. Available at: <https://www.hyperemesis.org/research/genetics-faq> (Accessed: 27 June 2025).
- [95] Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R. & Grosveld, F. (1999) The transcription factor GATA6 is essential for early extraembryonic development. *Development*, 126(9), pp. 723–732. Available at: doi:10.1242/dev.126.9.723.
- [96] Lynch, A.T., Phillips, N., Douglas, M., Dorgnach, M., Lin, I.H., Adamson, A.D., Darieva, Z., Whittle, J., Hanley, N.A., Bobola, N. & Birket, M.J. (2025) HAND1 level controls the specification of multipotent cardiac and extraembryonic progenitors from human pluripotent stem cells. *EMBO Journal*, 44(9), pp. 2541–2565. Available at: doi:10.1038/s44318-025-00409-0.
- [97] Fresch, R., Courtney, J., Brockway, H., Wilson, R.L. & Jones, H. (2023) HAND1 knockdown disrupts trophoblast global gene expression. *Physiological Reports*, 11(2), p. e15553. Available at: doi:10.14814/phy2.15553.
- [98] Chronopoulou, E., Koika, V., Tsiveriotis, K., Stefanidis, K., Kalogeropoulos, S., Georgopoulos, N., Adonakis, G. & Kaponis, A. (2022) Wnt4, Wnt6 and  $\beta$ -catenin expression in human placental tissue — is there a link with first trimester

- miscarriage? Results from a pilot study. *Reproductive Biology and Endocrinology*, 20(1), p. 51. Available at: doi:10.1186/s12958-022-00923-4.
- [99] Zhu Z., et al. (2024) Isl1 Identifies the Extraembryonic Mesodermal/Allantois Progenitors and Is Required for Placenta Morphogenesis and Vasculature Formation. *Advanced Science*. doi: 10.1002/advs.202400238.
- [100] Zhang, H., Qi, J., Guo, J., Wang, Y., Guan, Y., Fan, J., Sui, L., Xu, Y., Kong, L., Yan, B. & Kong, Y. (2021) Paired-like homeodomain transcription factor 2 affects endometrial cell function and embryo implantation through the Wnt/ $\beta$ -catenin pathway. *Cell Biology International*, 45(9), pp. 1957–1965. Available at: doi:10.1002/cbin.11636.
- [101] Shi, J.W. *et al.* (2020) Collagen at the maternal–fetal interface in human pregnancy. *International Journal of Biological Sciences*, 16(12), pp. 2220–2234. Available at: doi:10.7150/ijbs.45586.
- [102] Hiden, U., Majali-Martinez, A., Ghaffari-Tabrizi-Wizsy, N., Lang, U., Desoye, G. & Dieber-Rotheneder, M. (2016) Placental membrane-type metalloproteinases (MT-MMPs): key players in pregnancy. *Cell Adhesion & Migration*, 10(1–2), pp. 136–146. Available at: doi:10.1080/19336918.2015.1110671.
- [103] Klein, A.B. *et al.* (2023) Cross-species comparison of pregnancy-induced GDF15. *American Journal of Physiology — Endocrinology and Metabolism*, 325(4), pp. E303–E309. Available at: doi:10.1152/ajpendo.00134.2023.
- [104] Rodriguez R., et al. (2021) A low-cost mass spectrometry-based approach for quantifying purines in placental explants. *International Journal of Mass Spectrometry* 460:116490. doi: 10.1016/j.ijms.2020.116490.
- [105] Doughty, I.M., Glazier, J.D., Powell, T.L., Jansson, T. and Sibley, C.P. (1998) "Chloride transport across syncytiotrophoblast microvillous membrane of first

- trimester human placenta", *Pediatric Research*, 44(2), pp. 226–232. Available at: PMID: 9702919.
- [106] Riquelme, G. (2009) "Placental chloride channels: a review", *Placenta*. PMID: 19604577. Available at: <https://pubmed.ncbi.nlm.nih.gov/19604577>.
- [107] Barchetta, I., Ceccarelli, V., Cimini, F.A., Chiappetta, C., Bertocchini, L., Manca di Villahermosa, S., Porzio, O., Cavallo, M.G. and Baroni, M.G. (2022) "Neurotensin promotes inflammation and cell proliferation in visceral adipose tissue of obese subjects", *International Journal of Obesity*, 46(4), pp. 778–786. Available at: doi:10.1038/s41366-022-01125-8.
- [108] Joo, M., Sadikot, R.T., PGDS and PGD2 in immune response. (2012) *Mediators of Inflammation*, 2012, article ID 503128. Available at: doi:10.1155/2012/503128.
- [109] Shamas-Din, A., Kale, J., Leber, B. and Andrews, D.W. (2013) "Mechanisms of Action of Bcl-2 Family Proteins", *Cold Spring Harbor Perspectives in Biology*, 5(4), a008714. <https://doi.org/10.1101/cshperspect.a008714>.
- [110] Feng, Y., Wang, X., Lu, Z., Li, X., Liu, J. and Zhang, W. (2022) "Steroidogenic enzyme HSD3B2 in human reproductive tissues and its role in hormone synthesis", *Frontiers in Endocrinology*, 13, article 934515. Available at: doi:10.3389/fendo.2022.934515.
- [111] Brown, L.M., Rodriguez, A.C., Dey, S.K. and Wang, H. (2022) "Iroquois homeobox transcription factor 3 (IRX3) modulates endometrial receptivity and embryo implantation", *Reproductive Biology and Endocrinology*, 20(1), article 97. Available at: doi:10.1186/s12958-022-00961-8.
- [112] Li, D., Sakuma, R., Vakili, N.A., Mo, R., Puvindran, V., Deimling, S., Zhang, X., Hopyan, S. and Hui, C.C. (2014) "Formation of proximal and anterior limb skeleton requires early function of Irx3 and Irx5 and is negatively regulated by

- Shh signaling", *Developmental Cell*, 29(2), pp. 233–240. Available at: doi:10.1016/j.devcel.2014.03.004.
- [113] Kumari, S., Dash, R., Singh, R., Reddy, R.S., Ghosh, S. and Swarnakar, S. (2017) "Lysyl oxidase: its role in cancer progression and potential as a therapeutic target", *Critical Reviews in Oncology/Hematology*, 115, pp. 67–81. Available at: doi:10.1016/j.critrevonc.2017.05.003.
- [114] Lynch, A.T., Phillips, N., Douglas, M., Dorgnach, M., Lin, I.H., Adamson, A.D., Darieva, Z., Whittle, J., Hanley, N.A., Bobola, N. & Birket, M.J. (2025) HAND1 level controls the specification of multipotent cardiac and extraembryonic progenitors from human pluripotent stem cells. *EMBO Journal*, 44(9), pp. 2541–2565. Available at: doi:10.1038/s44318-025-00409-0.
- [115] Franco, D., Sedmera, D. & Lozano-Velasco, E. (2017) Multiple roles of Pitx2 in cardiac development and disease. *Journal of Cardiovascular Development and Disease*, 4(4), p. 16. Available at: doi:10.3390/jcdd4040016.
- [116] Zhang, H., Qi J., Guo, J., Wang, Y., Guan, Y., Fan, J., Sui, L., Xu, Y., Kong, L., Yan, B., Kong, Y. (2021) "Paired-like homeodomain transcription factor 2 affects endometrial cell function and embryo implantation through the Wnt/ $\beta$ -catenin pathway". *Cell Biol Int.* 45(9):1957-1965. Available at: doi: 10.1002/cbin.11636
- [117] Zhu, Z., Zou, Q., Wang, C., Li, D., Yang, Y., Xiao, Y., Jin, Y., Yan, J., Luo, L., Sun, Y. & Liang, X. (2024) Isl1 identifies the extraembryonic mesodermal/allantois progenitors and is required for placenta morphogenesis and vasculature formation. *Advanced Science*, 11(32), p. e2400238. Available at: doi:10.1002/advs.202400238
- [118] Wamaitha, S.E., del Valle, I., Cho, L.T., Wei, Y., Fogarty, N.M., Blakeley, P., Sherwood, R.I., Ji, H. & Niakan, K.K. (2015) Gata6 potently initiates

reprogramming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes & Development*, 29(12), pp. 1239–1255. Available at: doi:10.1101/gad.257071.114.

- [119] Zinani, OQH., Keseroğlu, K., Özbudak, EM. (2022) "Regulatory mechanisms ensuring coordinated expression of functionally related genes". *Trends Genet.* 38(1):73-81. Available at doi: 10.1016/j.tig.2021.07.008.
- [120] Winn, VD., Haimov-Kochman, R., Paquet, AC., Yang, YJ., Madhusudhan, MS., Gormley, M., Feng, KT., Bernlohr, DA., McDonagh, S., Pereira, L., Sali, A., Fisher, SJ. (2007) Gene expression profiling of the human maternal-fetal interface reveals dramatic changes between midgestation and term. *Endocrinology*. 148(3):1059-79. Available at doi:10.1210/en.2006-0683).
- [121] Knox K, Baker JC (2008) Genomic evolution of the placenta using co-option and duplication and divergence. *Genome Res* 18: 695–705.
- [122] Zhu, C., Jiang, Z., Bazer, FW., Johnson, GA., Burghardt, RC., Wu, G.. (2015) "Aquaporins in the female reproductive system of mammals". *Front Biosci* (Landmark Ed). 20(5):838-71. doi: 10.2741/4341.
- [123] Davis, A.P., Wiegers, T.C., Johnson, R.J., Sciaky, D., Wiegers, J. and Mattingly, C.J. (2023) "Comparative Toxicogenomics Database (CTD): update 2023", *Nucleic Acids Research*, 51(D1), pp. D1257–D1262. <https://doi.org/10.1093/nar/gkac833>
- [124] Vilotić, A., Nacka-Aleksić, M., Pirković, A., Bojić-Trbojević, Ž., Dekanski, D. and Jovanović Krivokuća, M. (2022) "IL-6 and IL-8: An Overview of Their Roles in Healthy and Pathological Pregnancies", *International Journal of Molecular Sciences*, 23(23), p. 14574. <https://doi.org/10.3390/ijms232314574>
- [125] Oishi, M., Inagaki, C., Takaori, S., Yajima, H. and Akazawa, Y. (1982) "Attenuation of triphasic blood pressure responses induced by shortening of

amino acid sequences of neurotensin”, *Neuropeptides*, 2(5), pp. 279–286.  
[https://doi.org/10.1016/0143-4179\(82\)90017-8](https://doi.org/10.1016/0143-4179(82)90017-8)

- [126] Osadchii, O.E. (2015) “Emerging role of neurotensin in regulation of the cardiovascular system”, *European Journal of Pharmacology*, 762, pp. 184–192.  
<https://doi.org/10.1016/j.ejphar.2015.05.025>
- [127] Bochkova, M.S., Zamorina, S.A., Timganova, V.P., Khramtsov, P.V. and Rayev, M.B. (2019) “The role of glycodelin in the regulation of the immune system in the context of developing pregnancy”, *Medical Immunology*, 21(4), pp. 603–616.  
<https://doi.org/10.15789/1563-0625-2019-4-603-616>
- [128] Knox, K., Baker, J.C. (2008) Genomic evolution of the placenta using co-option and duplication and divergence. *Genome Res* 18: 695–705.
- [129] Choi, G.B., Yim, Y.S., Wong, H., Kim, S., Kim, H., Kim, S.V., Hoeffler, C.A., Littman, D.R. and Huh, J.R. (2016) “The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring”, *Science*, 351(6276), pp. 933–939. <https://doi.org/10.1126/science.aad0314>
- [130] Moreau, P., Adrian-Cabestre, F., Menier, C., Guiard, V., Gourand, L., Dausset, J., Carosella, E.D. and Paul, P. (1999) “IL-10 selectively induces HLA-G expression in human trophoblasts and monocytes”, *International Immunology*, 11(5), pp. 803–811. <https://doi.org/10.1093/intimm/11.5.803>

## ДОДАТОК А

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ, В ЯКИХ  
ОПУБЛІКОВАНІ НАУКОВІ РЕЗУЛЬТАТИ ДИСЕРТАЦІЇ

Статті у наукових фахових виданнях

1. **Lykhenko, O.**, Frolova, A., & Obolenskaya, M. (2021) ‘Changes in the human placental transcriptome during the physiological course of pregnancy’, *Biopolym. Cell*, 37(1), pp. 73–82. Available at: doi:10.7124/bc.000A4D. (Scopus, Q4) (Особистий внесок здобувача – застосування розробленого пайплайну до даних експресії в здоровій плаценті першого, другого та третього триместрів, біологічна характеристика знайдених диференційно експресованих генів, написання статті та підготування її до друку).
2. **Lykhenko, O.**, & Obolenskaya, M. Yu. (2021) ‘The knowledge of chromosomal sex is important for large-scale analysis of gene expression’, *Dopov. Nac. akad. nauk Ukr.*, 2021(1), pp. 100–109. Available at: doi:10.15407/dopovidi2021.01.100. (Особистий внесок здобувача – визначення статі плоду за даними експресії, покращення якості пайплайну інтегративного аналізу завдяки даним про статі, написання статті та підготування її до друку).
3. **Lykhenko, O.**, Frolova, A., & Obolenskaya, M. (2021) ‘Consecutive integration of available microarray data for analysis of differential gene expression in human placenta’, *Biotechnologia Acta*, 14(1), pp. 38–45. Available at: doi:10.15407/biotech14.01.38. (Особистий внесок здобувача – програмування пайплайну обробки та інтеграції даних експресії, написання статті та підготування її до друку).
4. **Lykhenko, O.**, Frolova, A., & Obolenska, M. (2017) ‘Designing the database for microarray experiments metadata’, *2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF)*. IEEE. Available at:

doi:10.1109/YSF.2017.8126658. (Особистий внесок здобувача – розробка бази метаданих IGEA, написання статті та підготування її до друку).

5. **Lykhenko, O.**, Frolova, A., & Obolenskaya, M. (2017) ‘Creation of gene expression database on preeclampsia-affected human placenta’, *Biopolym. Cell*, 33(6), pp. 442–452. Available at: doi:10.7124/bc.000967. (Scopus, Q4) (Особистий внесок здобувача – відбір датасетів для дослідження, стандартизація та розширення метаданих, написання статті та підготування її до друку).

#### Тези наукових конференцій

1. Frolova A, **Lykhenko O.**, Bondarenko V., and Obolenskaya M. (2017) ‘Step by step to integration of gene expression profiles for preeclampsia-affected human placenta’, *China — Ukraine International Symposium on Innovation and Technology*, 30.10 — 03.11. 2017. Shandong Academy of sciences. Abstract book. pp. 1–8.
2. Зубенко, С.В., **Лихенко, О.К.**, Фролова, А.О. та Оболенська, М.Ю. (2021) ‘Comparing alternative pipelines for RNA-seq differential expression analysis between preeclamptic and control placental samples’. У: XV Всеукраїнська конференція молодих вчених з міжнародною участю, присвячена 30-річчю Незалежності України, 26 травня 2021, ІМБіГ НАН України, Київ.
3. **Lykhenko, O.**, Zhoha V., Lykhenko DK., Frolova A., Martsenyuk O., Obolenska M. (2019) «How smoking may decrease the risk of preeclampsia development». *Medical and Clinical Chemistry Scientific Journal*, I. Horbachevsky Ternopil National Medical University Institute of Pharmacology and Toxicology of NAMS of Ukraine, 3(80), vol. 21, p. 110.
4. Жалілова, Є.О., **Лихенко, О.К.** та Оболенська, М.Ю. (2023) ‘Топ-двадцять

диференційно експресованих генів у плаценті людини між другим і першим триместрами фізіологічної вагітності». У: «Біотехнологія ХХІ століття»: матеріали XVII Міжнародної науково-практичної конференції, 19 травня 2023, Київ: КПІ ім. Ігоря Сікорського, секція 2, с. 225–227.

5. Кукуруза, Є.О., **Лихенко, О.К.** та Оболенська, М.Ю. (2023) ‘Алгоритм пошуку тканинно-специфічних генів, продукти яких секретуються у кров, на прикладі плаценти людини’. У: «Біотехнологія ХХІ століття»: матеріали XVII Міжнародної науково-практичної конференції, 19 травня 2023, Київ: КПІ ім. Ігоря Сікорського, секція 2, с. 228–232.
6. Жалілова, Є.О., **Лихенко, О.К.** та Оболенська, М.Ю. (2024) ‘Pathway enrichment analysis in the human placenta for the period between the second and first trimesters of physiological pregnancy’. Тези Міжнародної конференції з нейронаук та наукових читань, присвячених вісцеральній фізіології та патофізіології, 19–21 листопада 2024, Київ. Фізіол. журн., 70(5) дод., с. 33.
7. Martynenko, V.V., Zhalilova, Y.O., **Lyhenko, O.K.** & Obolenskaya, M.Yu. (2024) ‘Top 20 differentially expressed genes across trimesters of normal pregnancy’, in Proceedings of the International Scientific Conference BioGENext, Kyiv, Ukraine, 17–20 September 2024. Biopolymers and Cell, 40(3), p. 210. doi: 10.7124/bc.000AEA.

## ДОДАТОК Б

### РОЗРОБКА СТРУКТУРИ БАЗИ ДАНИХ

Перша ідея для структури бази даних — завантажити з ArrayExpress таблиці з даними про експерименти в цілому, про платформи мікрочипів та про кожен окремий зразок і відтворити ці таблиці (Experiment, Microarray, Sample) в нашій спеціалізованій базі даних. Кожний рядок відповідав би експерименту, мікрочипу і зразку, а кожна колонка — атрибуту (характеристиці зразка). Такий варіант структури бази даних наведений на рис. 1.

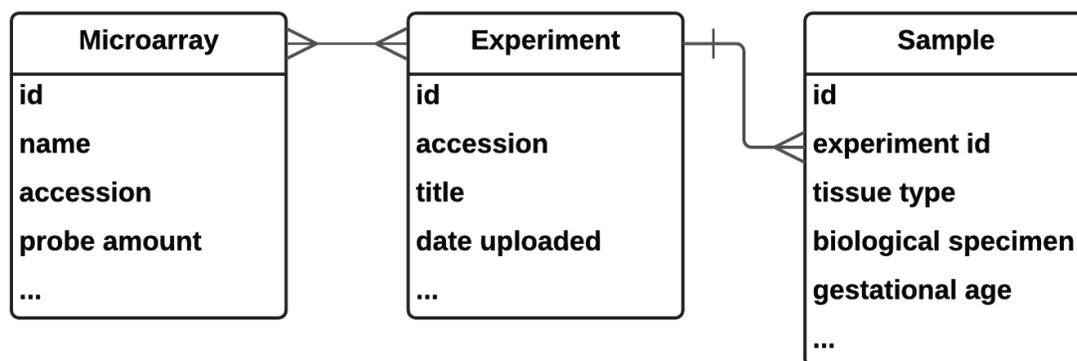


Рис. 1. Початкова структура бази даних

Ця структура проста, але вона вимагатиме додавати нову колонку кожного разу, коли з'являтиметься нова характеристика в кожній з таблиць. У той час, як експерименти та мікрочипи в ArrayExpress здебільшого мають стандартні набори характеристик, біологічні зразки з різних експериментів можуть мати різні характеристики, і тому кожна нова характеристика зразка вимагатиме модифікації структури бази даних. Таких модифікацій варто уникати, адже вони можуть призвести до несумісності різних версій однієї і тієї ж бази даних. До того ж, оскільки не всі зразки матимуть більшість характеристик, то в таблиці Samples

буде багато пустих клітинок, які займатимуть простір на жорсткому диску комп'ютера.

Вирішити цю проблему можна шляхом заміни численних колонок-характеристик однією колонкою складнішої структури. СУБД (система управління базами даних) Postgres має поле HStore, яке по суті є асоціативним масивом, тобто набором пар рядків типу “ключ-значення”, де кожен ключ з’являється в наборі не більше одного разу. Заміна атрибутів єдиним полем HStore, назовемо його “data”, надає можливість додавати інформацію про нові характеристики не змінюючи структури бази даних (рис. 2).

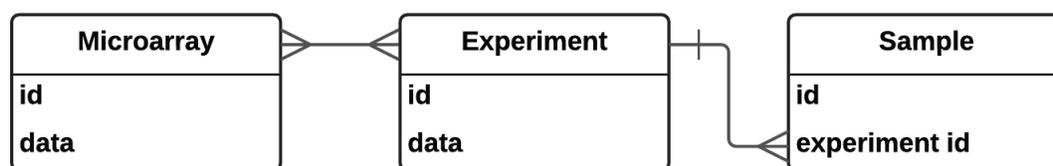


Рис. 2. Майже безструктурна база даних; data — поля типу HStore

У такій базі даних, якщо ми завантажимо з ArrayExpress зразок з атрибутом “age” зі значенням “37 weeks” та атрибут “diagnosis” зі значенням “pre-eclampsia”, метадані будуть збережені в колонці “data”, значення якої для даного зразка дорівнюватиме “age”:”37 weeks”, “diagnosis”:”pre-eclampsia”, замість того, щоб мати окремі колонки “age”, “diagnosis” та скоріш за все багато пустих колонок для характеристик, які відсутні для одного зразка, але присутні для інших.

Крім того, коли ми завантажили дані з ArrayExpress і почали стандартизувати назви і значення характеристик зразків, для нас стало доцільним зберігати первинні завантажені з ArrayExpress назви та значення разом зі

стандартизованими. Через це ми більше не можемо зберігати атрибути зразків в полі “data” типу HStore, оскільки тепер ми хочемо зберігати не два рядки, стандартизовані нову назву і нове значення, а чотири рядки для кожного атрибута, додаючи первинні назву і значення. Щоб вирішити цю проблему, ми замінили колонку “data” колонкою типу зовнішній ключ (foreignkey), значення якої вказуватимуть на рядки нової таблиці SampleAttribute, яка міститиме всі атрибути всіх зразків. В термінології баз даних такий зв’язок між таблицями зветься *один-до-багатьох*. Таблиця SampleAttribute у свою чергу містить зовнішні ключі до таблиць SampleAttributeName та SampleAttributeValue, що містять старі й нові імена та значення атрибутів: (рис. 3)

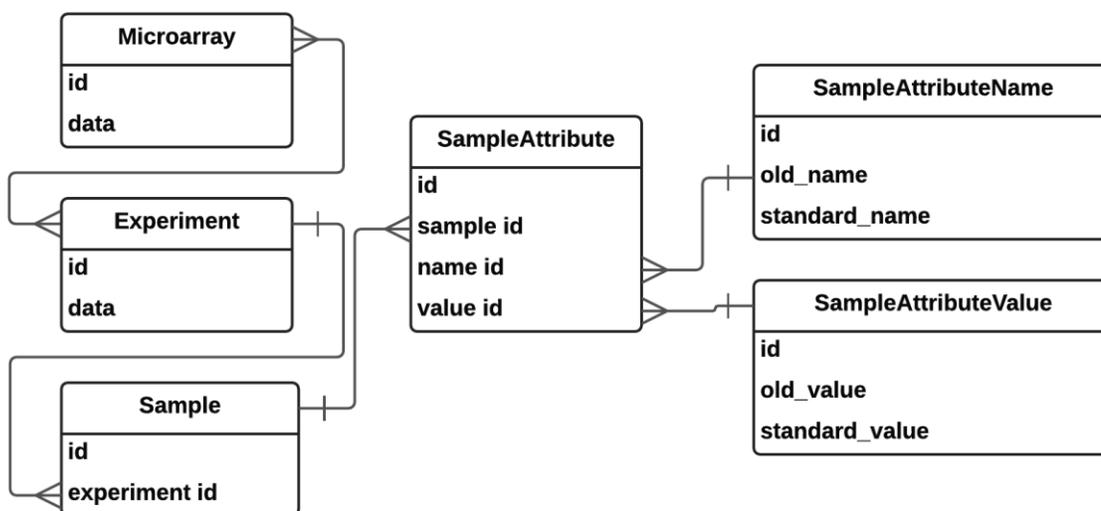


Рис. 3. Структура бази даних зі збереженими початковими іменами та значеннями атрибутів

У таблиці з такою структурою спостерігається значне дублювання даних, зокрема значень полів “standard\_name” та “standard\_value”, оскільки одному “old\_value” може відповідати кілька “standard\_value” залежно від контексту експерименту. Наприклад, початкове “old\_value” “placenta” може відповідати

“standard\_value” “Chorion”, якщо в тексті статті вказується саме ця тканина, або ж “Placenta” якщо не вказується назва частини матково-плацентарного комплексу.

Щоб позбутися дублювання даних, було б непогано, якби стандартні назви та значення описувалися в базі даних один раз, і були присутніми у якості посилання там, де вони необхідні для опису зразків, а початкові імена та значення атрибутів були б безпосередньо пов’язані з конкретним атрибутом зразка, як на рис. 4.

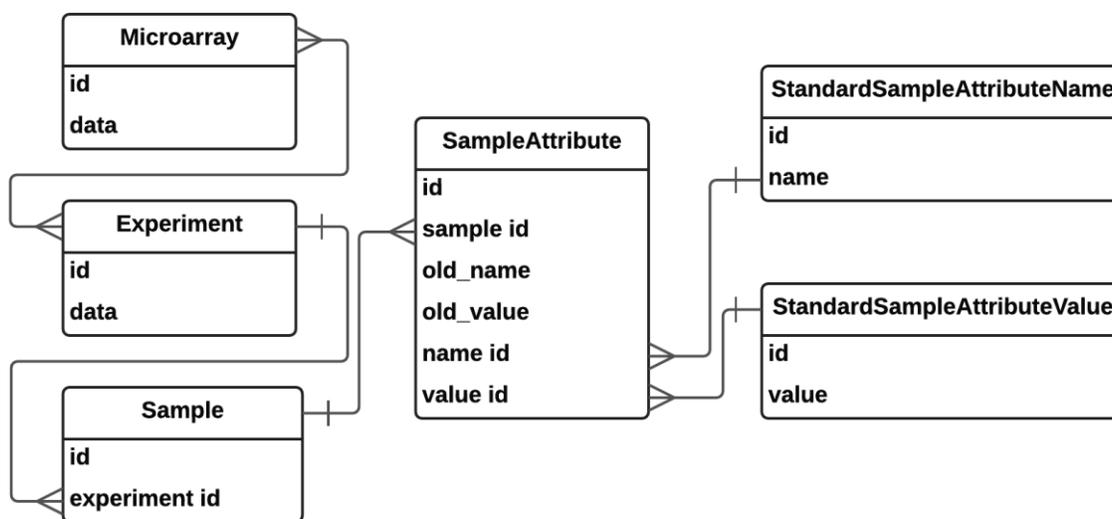


Рис. 4. Модифікована структура, бази даних, у якій дані не дублюються

Протягом стандартизації термінології виявилось, що не всі терміни мають відповідний термін у MeSH. В окремих випадках нам довелося звернутися до онтології Experimental Factor Ontology (EFO) та запроваджувати власні терміни. Для збереження додаткової інформації про StandardSampleAttributeName та StandardSampleAttributeValue ми додали поле “additional\_info” типу HStore до цих таблиць. Щоб вказати, що деякі стандартні імена є спорідненими (синоніми або ширше-вужче поняття), ми додали рефлексивне (яке посилається саме на себе) поле Synonyms типу багато-до-багатьох таблиць StandardSampleAttributeName та

StandardSampleAttributeValue. Кінець кінцем, ми додали зв'язок один-до-багатьох між StandardSampleAttributeName та StandardSampleAttributeValue, щоб показати, які значення відповідають яким назвам (рис. 5).

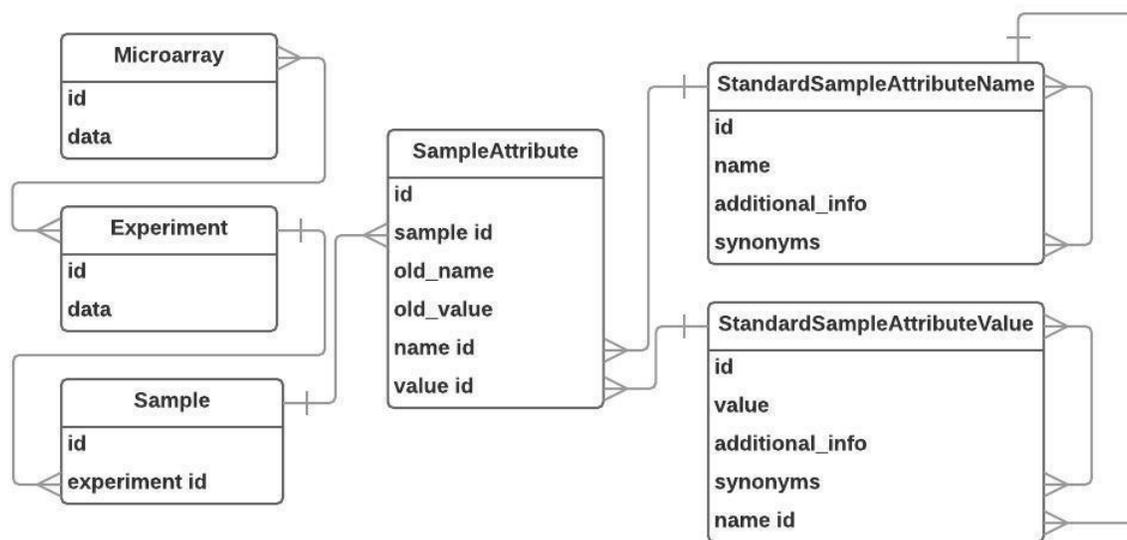


Рис 5. Схематичне зображення бази даних

1) *Примітка 1:* Значення атрибутів зразків можуть бути якісними (наприклад, триместр, стать) і кількісними (вік гестації, маса плоду). Для кількісних атрибутів, значення зберігається в полі “old\_value”, а в “standard\_value” зберігаються одиниці вимірювання (“numberinweeks”, “massingrams”). Таким чином, атрибут зразка з назвою “Gestation Age” матиме “standard\_value” зі значенням “Numberinweeks” та “old\_value” зі значеннями 42.

2) *Примітка 2:* Деякі імена та значення атрибутів були вилучені з процесу стандартизації та поставлені у відповідність зі стандартним іменем або значенням “(excluded)” через те, що вони були або нерелевантними або інформація була перенесена до інших атрибутів (“UnknownSex”, “Other”, “cal id”, “passage”, “matching”, “N/A”). Вони були позначені як “<empty>”.

**ДОДАТОК В**  
**ДИФЕРЕНЦІЙНО ЕКСПРЕСОВАНІ ГЕНИ МІЖ І-ІІ ТРИМЕСТРАМИ ЗА**  
**КЛАСТЕРАМИ**

SYMBOL	ENTREZID	GENENAME	logFC	adj.P.Val (FDR)	cluster
ST6GALNAC1	55808	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 1	-1.11	2.11E-02	1
PAG1	55824	phosphoprotein membrane anchor with glycosphingolipid microdomains 1	1.12	1.20E-03	1
GIMAP2	26157	GTPase, IMAP family member 2	1.22	3.88E-03	1
CCL2	6347	C-C motif chemokine ligand 2	1.65	3.87E-04	1
CCL13	6357	C-C motif chemokine ligand 13	1.52	1.59E-04	1
SRGN	5552	serglycin	1.02	1.90E-02	1
CD93	22918	CD93 molecule	1.02	9.89E-04	1
CCR7	1236	C-C motif chemokine receptor 7	-1.27	2.06E-03	1
BST2	684	bone marrow stromal cell antigen 2	1.01	3.26E-05	1
RNF128	79589	ring finger protein 128	1.27	2.94E-02	1
LYVE1	10894	lymphatic vessel endothelial hyaluronan receptor 1	1.84	4.71E-03	1
NPL	80896	N-acetylneuraminase pyruvate lyase	1.04	1.58E-02	1
SLC19A3	80704	solute carrier family 19 member 3	1.24	4.30E-03	1
CCL21	6366	C-C motif chemokine ligand 21	-1.24	1.17E-02	1
PI15	51050	peptidase inhibitor 15	-1.79	8.26E-03	1
CYP2C9	1559	cytochrome P450 family 2 subfamily C member 9	-1.41	2.89E-03	1
EPCAM	4072	epithelial cell adhesion molecule	-3.08	1.55E-03	1
F13A1	2162	coagulation factor XIII A chain	1.29	3.89E-02	1
ALDH8A1	64577	aldehyde dehydrogenase 8 family member A1	-1.70	8.67E-03	1
BCL2A1	597	BCL2 related protein A1	1.04	3.64E-03	1
REN	5972	renin	-1.24	6.17E-04	1
TIMD4	91937	T cell immunoglobulin and mucin domain containing 4	1.33	1.26E-03	1
ESAM	90952	endothelial cell adhesion molecule	1.12	5.63E-05	1
TMEM45B	120224	transmembrane protein 45B	-1.47	2.55E-03	1
GLT1D1	144423	glycosyltransferase 1 domain containing 1	-1.42	2.72E-03	1
HGD	3081	homogentisate 1,2-dioxygenase	-1.18	7.38E-03	1
ATP8B4	79895	ATPase phospholipid transporting 8B4 (putative)	1.23	2.70E-04	1
CXCL13	10563	C-X-C motif chemokine ligand 13	-2.18	3.95E-03	1
SLCO2B1	11309	solute carrier organic anion transporter family member 2B1	1.03	1.93E-05	1
PDZK1IP1	10158	PDZK1 interacting protein 1	-1.52	3.26E-02	1
HPGDS	27306	hematopoietic prostaglandin D synthase	1.13	3.11E-02	1

PPBP	5473	pro-platelet basic protein	2.00	5.38E-04	1
PF4	5196	platelet factor 4	1.98	3.64E-04	1
DEFB1	1672	defensin beta 1	-3.23	1.49E-03	1
FOLR2	2350	folate receptor beta	1.02	1.78E-03	1
CLDN10	9071	claudin 10	-1.62	3.01E-03	1
MS4A7	58475	membrane spanning 4-domains A7	1.06	8.69E-03	1
RNASE6	6039	ribonuclease A family member k6	1.13	3.99E-03	1
C3AR1	719	complement C3a receptor 1	1.19	2.19E-03	1
CXCL8	3576	C-X-C motif chemokine ligand 8	1.74	1.40E-02	1
C1QB	713	complement C1q B chain	1.37	3.62E-03	1
C7	730	complement C7	2.06	8.76E-06	1
FGF13	2258	fibroblast growth factor 13	1.24	8.67E-04	1
IRX3	79191	iroquois homeobox 3	-2.35	7.38E-03	1
HBB	3043	hemoglobin subunit beta	1.06	3.89E-02	1
RNASE4	6038	ribonuclease A family member 4	-1.05	3.63E-02	1
MS4A4A	51338	membrane spanning 4-domains A4A	1.62	1.93E-03	1
HLA-DQA1	3117	major histocompatibility complex, class II, DQ alpha 1	-1.25	9.35E-03	1
SLPI	6590	secretory leukocyte peptidase inhibitor	-2.19	2.64E-03	1
LVRN	206338	laeverin	1.25	2.83E-02	1
FCGR2B	2213	Fc fragment of IgG receptor IIb	2.97	1.40E-05	1
CD163	9332	CD163 molecule	1.29	3.20E-03	1
MPEG1	219972	macrophage expressed 1	1.29	8.27E-05	1
HSD11B1	3290	hydroxysteroid 11-beta dehydrogenase 1	2.66	4.94E-06	1
NCF2	4688	neutrophil cytosolic factor 2	1.10	5.08E-03	1
MNDA	4332	myeloid cell nuclear differentiation antigen	1.39	1.06E-03	1
S100A8	6279	S100 calcium binding protein A8	1.87	9.60E-03	1
S100A9	6280	S100 calcium binding protein A9	1.69	1.04E-04	1
TMIGD3	57413	transmembrane and immunoglobulin domain containing 3	1.46	1.31E-03	1
TSPAN1	10103	tetraspanin 1	-2.04	5.52E-03	1
WFDC2	10406	WAP four-disulfide core domain 2	-1.87	5.85E-03	1
LCN2	3934	lipocalin 2	-1.99	8.10E-03	1
C1QC	714	complement C1q C chain	1.13	1.19E-02	1
C1QA	712	complement C1q A chain	1.33	1.08E-03	1
VSIG4	11326	V-set and immunoglobulin domain containing 4	1.02	2.37E-02	1
CARD16	114769	caspase recruitment domain family member 16	1.18	4.98E-04	1
TNFSF13B	10673	TNF superfamily member 13b	1.37	2.84E-04	1
LST1	7940	leukocyte specific transcript 1	1.04	7.16E-04	1
GPR183	1880	G protein-coupled receptor 183	1.54	2.11E-05	1
GPR34	2857	G protein-coupled receptor 34	1.55	2.49E-02	1
CYBB	1536	cytochrome b-245 beta chain	1.47	5.43E-04	1
TLR7	51284	toll like receptor 7	1.14	4.90E-04	1
PFKP	5214	phosphofructokinase, platelet	-1.34	2.83E-04	1
IL33	90865	interleukin 33	1.13	6.82E-03	1
GPNMB	10457	glycoprotein nmb	1.57	1.08E-03	1
LAIR1	3903	leukocyte associated immunoglobulin like receptor 1	1.34	1.80E-05	1

BPGM	669	bisphosphoglycerate mutase	1.11	3.78E-02	1
LGMN	5641	legumain	1.04	3.61E-05	1
TSPAN8	7103	tetraspanin 8	-3.13	2.19E-03	1
METTL7B	196410	methyltransferase like 7B	1.32	3.51E-05	1
DNASE1L3	1776	deoxyribonuclease 1 like 3	1.11	3.22E-03	1
CCL8	6355	C-C motif chemokine ligand 8	1.14	1.09E-03	1
CLDN3	1365	claudin 3	-1.71	1.07E-03	1
AREG	374	amphiregulin	1.58	1.18E-02	1
CXCL1	2919	C-X-C motif chemokine ligand 1	1.53	1.18E-02	1
ADHFE1	137872	alcohol dehydrogenase iron containing 1	1.02	7.73E-03	1
GATM	2628	glycine amidinotransferase	1.08	5.86E-04	1
NCF4	4689	neutrophil cytosolic factor 4	1.03	1.85E-04	1
RNASE1	6035	ribonuclease A family member 1, pancreatic	1.40	1.82E-03	1
IFI30	10437	IFI30 lysosomal thiol reductase	1.03	2.33E-02	1
AGR2	10551	anterior gradient 2, protein disulphide isomerase family member	-1.82	5.84E-03	1
MS4A6A	64231	membrane spanning 4-domains A6A	1.21	1.72E-03	1
CD36	948	CD36 molecule	1.72	3.99E-03	1
LILRB5	10990	leukocyte immunoglobulin like receptor B5	1.27	3.93E-06	1
TMEM176B	28959	transmembrane protein 176B	1.73	1.60E-04	1
AOC1	26	amine oxidase copper containing 1	1.65	1.04E-02	1
MUC20	200958	mucin 20, cell surface associated	-1.58	1.58E-02	1
CFB	629	complement factor B	1.04	4.01E-02	1
TMEM176A	55365	transmembrane protein 176A	1.06	3.74E-03	1
CXCL2	2920	C-X-C motif chemokine ligand 2	1.37	1.18E-02	1
GLB1L2	89944	galactosidase beta 1 like 2	-1.00	4.25E-02	1
HCK	3055	HCK proto-oncogene, Src family tyrosine kinase	1.23	6.40E-05	1
HSD3B2	3284	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2	2.06	4.94E-06	1
EMX2	2018	empty spiracles homeobox 2	-1.89	3.78E-03	1
RNASE4	6038	ribonuclease A family member 4	-1.05	3.63E-02	1
MRC1	4360	mannose receptor C-type 1	1.17	1.86E-02	1
PECAM1	5175	platelet and endothelial cell adhesion molecule 1	1.46	7.70E-06	1
TPPP3	51673	tubulin polymerization promoting protein family member 3	1.22	2.98E-06	1
FPR1	2357	formyl peptide receptor 1	1.43	2.17E-04	1
RNASE4	6038	ribonuclease A family member 4	-1.05	3.63E-02	1
GIMAP6	474344	GTPase, IMAP family member 6	1.03	1.27E-02	1
ALOX5AP	241	arachidonate 5-lipoxygenase activating protein	1.05	4.98E-02	1
HNF1B	6928	HNF1 homeobox B	-1.01	8.67E-04	1
MUC1	4582	mucin 1, cell surface associated	-1.06	1.07E-02	1
HSD17B2	3294	hydroxysteroid 17-beta dehydrogenase 2	1.77	7.80E-06	2
GADD45B	4616	growth arrest and DNA damage inducible beta	1.01	8.27E-05	2
PLAT	5327	plasminogen activator, tissue type	-1.34	4.32E-02	2
COMP	1311	cartilage oligomeric matrix protein	-1.76	2.41E-03	2
TFPI2	7980	tissue factor pathway inhibitor 2	1.46	4.23E-02	2
ISL1	3670	ISL LIM homeobox 1	1.35	6.53E-03	2
LOX	4015	lysyl oxidase	-1.37	4.26E-02	2

SEMA3G	56920	semaphorin 3G	1.07	6.25E-04	2
MAGEA10	4109	MAGE family member A10	1.06	4.85E-04	2
SCGB1D2	10647	secretoglobin family 1D member 2	-2.00	5.47E-03	2
SCGB2A1	4246	secretoglobin family 2A member 1	-3.20	1.86E-03	2
GNG11	2791	G protein subunit gamma 11	1.33	2.98E-06	2
HOXD10	3236	homeobox D10	-1.07	1.62E-03	2
ISLR	3671	immunoglobulin superfamily containing leucine rich repeat	-1.20	3.35E-02	2
NTS	4922	neurotensin	1.43	1.06E-02	2
ENPP2	5168	ectonucleotide pyrophosphatase/phosphodiesterase 2	1.07	2.13E-04	2
CCN3	4856	cellular communication network factor 3	-1.56	8.20E-04	2
FOXF2	2295	forkhead box F2	1.03	1.40E-03	2
EXOC6	54536	exocyst complex component 6	1.20	8.17E-06	2
MYO5C	55930	myosin VC	-1.14	1.32E-02	2
CILP	8483	cartilage intermediate layer protein	-1.37	4.03E-03	2
MEOX2	4223	mesenchyme homeobox 2	1.81	1.95E-03	2
SLC1A1	6505	solute carrier family 1 member 1	-2.19	3.98E-03	2
PRKX	5613	protein kinase X-linked	-1.23	7.57E-04	2
CNTN3	5067	contactin 3	-1.19	5.28E-03	2
GRB14	2888	growth factor receptor bound protein 14	1.50	2.98E-06	2
GUCY1B1	2983	guanylate cyclase 1 soluble subunit beta 1	1.06	3.25E-04	2
CP	1356	ceruloplasmin	-2.03	2.24E-02	2
PRKAR2B	5577	protein kinase cAMP-dependent type II regulatory subunit beta	1.34	2.63E-05	2
ERP27	121506	endoplasmic reticulum protein 27	-1.35	2.39E-03	2
SFRP2	6423	secreted frizzled related protein 2	-1.33	4.52E-03	2
HHEX	3087	hematopoietically expressed homeobox	1.09	5.79E-05	2
ADAMTS18	170692	ADAM metalloproteinase with thrombospondin type 1 motif 18	1.09	4.50E-04	2
PTPRR	5801	protein tyrosine phosphatase receptor type R	-1.83	1.35E-02	2
ABI3BP	25890	ABI family member 3 binding protein	1.22	2.21E-02	2
DEPTOR	64798	DEP domain containing MTOR interacting protein	-1.58	5.02E-03	2
ACSM3	6296	acyl-CoA synthetase medium chain family member 3	-1.41	4.27E-03	2
PDPN	10630	podoplanin	1.13	5.99E-05	2
CLEC3B	7123	C-type lectin domain family 3 member B	1.42	2.97E-02	2
EMCN	51705	endomucin	-1.25	6.33E-03	2
GEM	2669	GTP binding protein overexpressed in skeletal muscle	-1.22	1.07E-02	2
ARHGAP42	143872	Rho GTPase activating protein 42	1.50	6.16E-06	2
JAM3	83700	junctional adhesion molecule 3	1.09	4.20E-05	2
IGFBP6	3489	insulin like growth factor binding protein 6	-1.46	7.77E-03	2
SOX7	83595	SRY-box transcription factor 7	1.13	1.36E-04	2
FRMD3	257019	FERM domain containing 3	1.22	1.59E-04	2
ADRB2	154	adrenoceptor beta 2	1.23	3.87E-04	2
VAMP5	10791	vesicle associated membrane protein 5	1.04	1.17E-03	2

NPFFR2	10886	neuropeptide FF receptor 2	1.37	1.37E-03	2
AQP1	358	aquaporin 1 (Colton blood group)	-1.12	8.27E-05	2
MZT2A	653784	mitotic spindle organizing protein 2A	-1.07	6.25E-06	2
PRG2	5553	proteoglycan 2, pro eosinophil major basic protein	1.34	1.59E-02	2
RPRM	56475	reprimin, TP53 dependent G2 arrest mediator homolog	-1.15	1.51E-03	2
CLU	1191	clusterin	-1.44	8.08E-03	2
KANK3	256949	KN motif and ankyrin repeat domains 3	1.23	6.37E-06	2
TCEAL7	56849	transcription elongation factor A like 7	-1.21	6.82E-03	2
PAPPA	5069	pappalysin 1	1.72	1.69E-03	2
GAST	2520	gastrin	-1.30	4.08E-03	2
ALDH1A3	220	aldehyde dehydrogenase 1 family member A3	-1.25	4.25E-02	2
FLRT2	23768	fibronectin leucine rich transmembrane protein 2	-1.42	3.82E-04	2
PLS1	5357	plastin 1	-1.81	1.44E-04	2
SFN	2810	stratifin	-1.79	1.15E-04	2
COL9A3	1299	collagen type IX alpha 3 chain	-1.37	4.92E-03	2
ANKRD55	79722	ankyrin repeat domain 55	-1.30	6.28E-03	2
CORO2A	7464	coronin 2A	-1.19	6.25E-06	2
ECM2	1842	extracellular matrix protein 2	-1.23	2.60E-02	2
DNER	92737	delta/notch like EGF repeat containing	-2.49	1.32E-02	2
SORBS2	8470	sorbin and SH3 domain containing 2	1.11	4.94E-06	2
MAP1LC3C	440738	microtubule associated protein 1 light chain 3 gamma	1.28	2.98E-06	2
PLAC1	10761	placenta enriched 1	1.79	1.95E-03	2
MAGEA4	4103	MAGE family member A4	1.38	1.62E-03	2
THBS2	7058	thrombospondin 2	-1.54	2.27E-02	2
ADGRG6	57211	adhesion G protein-coupled receptor G6	1.04	4.63E-02	2
DPT	1805	dermatopontin	1.50	1.34E-04	2
TCF21	6943	transcription factor 21	1.75	5.82E-05	2
MOXD1	26002	monooxygenase DBH like 1	-1.29	5.02E-03	2
CADM3	57863	cell adhesion molecule 3	1.18	1.89E-05	2
FAM162B	221303	family with sequence similarity 162 member B	1.57	5.05E-05	2
ADRB1	153	adrenoceptor beta 1	1.00	6.11E-03	2
KIAA1324	57535	KIAA1324	-1.08	4.77E-03	2
ABCA4	24	ATP binding cassette subfamily A member 4	-1.97	3.74E-03	2
BMP5	653	bone morphogenetic protein 5	1.96	2.11E-05	2
GADD45A	1647	growth arrest and DNA damage inducible alpha	-1.38	6.04E-03	2
PLA2G5	5322	phospholipase A2 group V	1.36	2.95E-05	2
GPC6	10082	glypican 6	-1.83	4.23E-05	2
LPAR6	10161	lysophosphatidic acid receptor 6	1.19	6.83E-04	2
NRCAM	4897	neuronal cell adhesion molecule	1.02	6.25E-06	2
HBE1	3046	hemoglobin subunit epsilon 1	-2.09	4.67E-02	2
COL2A1	1280	collagen type II alpha 1 chain	-1.39	2.37E-02	2
CYS1	192668	cystin 1	-1.87	1.14E-02	2
GLIS3	169792	GLIS family zinc finger 3	-1.15	4.05E-03	2
SEMA5A	9037	semaphorin 5A	-1.09	4.20E-04	2
LAMB3	3914	laminin subunit beta 3	-1.49	3.35E-02	2

GPC3	2719	glypican 3	1.02	1.77E-02	2
AFP	174	alpha fetoprotein	-2.38	1.69E-02	2
CXCL12	6387	C-X-C motif chemokine ligand 12	-1.17	2.44E-03	2
CCK	885	cholecystokinin	1.19	4.16E-02	2
JAM2	58494	junctional adhesion molecule 2	1.56	7.29E-06	2
SVEP1	79987	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	1.09	3.17E-06	2
IL6	3569	interleukin 6	1.69	3.36E-03	2
CPE	1363	carboxypeptidase E	-1.49	6.65E-03	2
CHN1	1123	chimerin 1	1.07	4.17E-03	2
SCN7A	6332	sodium voltage-gated channel alpha subunit 7	1.31	1.04E-04	2
ACAN	176	aggrecan	-1.04	3.76E-03	2
LHFPL3	375612	LHFPL tetraspan subfamily member 3	-1.51	1.29E-02	2
ATP2C2	9914	ATPase secretory pathway Ca <sup>2+</sup> transporting 2	-1.44	1.98E-03	2
SERPINB2	5055	serpin family B member 2	1.78	1.91E-02	2
COLEC11	78989	collectin subfamily member 11	-1.09	1.81E-05	2
OSR2	116039	odd-skipped related transcription factor 2	-1.25	3.09E-02	2
NECTIN3	25945	nectin cell adhesion molecule 3	1.48	6.29E-03	2
AGTR1	185	angiotensin II receptor type 1	1.31	1.39E-02	2
CD200	4345	CD200 molecule	1.13	3.01E-03	2
HEPH	9843	hephaestin	-1.65	1.14E-02	2
ADM	133	adrenomedullin	1.26	2.01E-02	2
BDKRB2	624	bradykinin receptor B2	-1.31	4.64E-03	2
TMEM100	55273	transmembrane protein 100	2.15	1.05E-04	2
HSD17B1	3292	hydroxysteroid 17-beta dehydrogenase 1	1.38	1.97E-02	2
APLNR	187	apelin receptor	1.23	1.16E-03	2
TMPRSS4	56649	transmembrane serine protease 4	-1.25	2.20E-03	2
NRN1	51299	neuritin 1	1.34	1.54E-03	2
TAC3	6866	tachykinin precursor 3	1.22	1.48E-02	2
SEMA3B	7869	semaphorin 3B	1.37	7.30E-03	2
PDK4	5166	pyruvate dehydrogenase kinase 4	1.55	1.06E-05	3
LIPG	9388	lipase G, endothelial type	1.39	8.67E-03	3
LPL	4023	lipoprotein lipase	1.27	7.49E-05	3
MT1F	4494	metallothionein 1F	-1.95	1.99E-04	3
C4BPA	722	complement component 4 binding protein alpha	-2.20	1.87E-03	3
MT1M	4499	metallothionein 1M	-1.92	5.66E-03	3
GPX3	2878	glutathione peroxidase 3	-1.73	3.27E-03	3
MT1X	4501	metallothionein 1X	-1.47	7.68E-04	3
MT1G	4495	metallothionein 1G	-2.60	2.90E-04	3
PAEP	5047	progesterone associated endometrial protein	-4.43	5.54E-04	3
EBI3	10148	Epstein-Barr virus induced 3	1.89	1.06E-02	4
IL12RB2	3595	interleukin 12 receptor subunit beta 2	1.25	2.88E-04	4
CSHL1	1444	chorionic somatomammotropin hormone like 1	2.25	1.26E-02	4
LEP	3952	leptin	-2.68	2.53E-03	4
GH1	2688	growth hormone 1	1.94	1.23E-02	4
CSH1	1442	chorionic somatomammotropin hormone 1	1.64	3.38E-02	4
CSF2RB	1439	colony stimulating factor 2 receptor beta common subunit	1.44	6.39E-03	4



**ДОДАТОК Г**  
**ДИФЕРЕНЦІЙНО ЕКСПРЕСОВАНИ ГЕНИ МІЖ II-III ТРИМЕСТРАМИ**  
**ЗА КЛАСТЕРАМИ**

<b>SYMBOL</b>	<b>ENTREZI D</b>	<b>GENENAME</b>	<b>logFC</b>	<b>adj.P.Val (FDR)</b>	<b>cluster</b>
CEACAM1	634	CEA cell adhesion molecule 1	1.94	4.08E-22	1
CETP	1071	cholesterol ester transfer protein	2.16	1.51E-14	1
АРОН	350	apolipoprotein H	-1.21	9.07E-08	1
IL2RB	3560	interleukin 2 receptor subunit beta	-1.63	7.16E-06	1
PROCR	10544	protein C receptor	1.04	3.09E-06	1
TIMP1	7076	TIMP metalloproteinase inhibitor 1	-1.15	1.31E-04	1
LHB	3972	luteinizing hormone subunit beta	1.42	3.86E-05	1
EBI3	10148	Epstein-Barr virus induced 3	1.98	7.99E-09	1
CCL2	6347	C-C motif chemokine ligand 2	-1.27	1.34E-04	1
CCL13	6357	C-C motif chemokine ligand 13	-1.29	1.22E-05	1
HLF	3131	HLF transcription factor, PAR bZIP family member	1.50	9.54E-10	1
VTN	7448	vitronectin	-2.21	2.83E-20	1
CXCL6	6372	C-X-C motif chemokine ligand 6	-1.33	3.27E-06	1
АРОС3	345	apolipoprotein C3	-1.07	4.77E-07	1
CHPT1	56994	choline phosphotransferase 1	-1.04	9.09E-13	1
REEP6	92840	receptor accessory protein 6	-1.04	1.88E-13	1
PLEK	5341	pleckstrin	-1.08	4.44E-08	1
RARRES1	5918	retinoic acid receptor responder 1	-1.21	5.39E-12	1
GDA	9615	guanine deaminase	1.09	2.75E-04	1
TGFB3	7043	transforming growth factor beta 3	-1.45	2.81E-09	1
AHR	196	aryl hydrocarbon receptor	1.35	3.49E-06	1
RAMP3	10268	receptor activity modifying protein 3	2.29	2.85E-19	1
TNFAIP6	7130	TNF alpha induced protein 6	-1.23	1.27E-05	1
MAGEA10	4109	MAGE family member A10	2.41	3.17E-18	1
TREM1	54210	triggering receptor expressed on myeloid cells 1	1.12	5.47E-05	1
C3	718	complement C3	-1.24	5.19E-05	1
GNGT1	2792	G protein subunit gamma transducin 1	1.14	7.84E-03	1
LIF	3976	LIF interleukin 6 family cytokine	-1.42	1.55E-09	1
АРОЕ	348	apolipoprotein E	-1.03	4.54E-08	1
EMILIN2	84034	elastin microfibril interfacier 2	-1.36	8.47E-10	1
CHI3L1	1116	chitinase 3 like 1	-1.34	5.27E-08	1
NTS	4922	neurotensin	-3.40	3.40E-35	1
FABP4	2167	fatty acid binding protein 4	5.49	1.42E-21	1
ADAMDEC1	27299	ADAM like decysin 1	1.61	2.27E-18	1
CD180	4064	CD180 molecule	-1.60	5.85E-19	1
FST	10468	follicle-stimulating hormone	-1.01	2.67E-07	1
NPL	80896	N-acetylneuraminidase	-1.23	8.60E-16	1
SCARB1	949	scavenger receptor class B member 1	1.80	5.22E-10	1

GNLY	10578	granulysin	-1.39	5.00E-03	1
KYNU	8942	kynureninase	1.03	1.37E-05	1
CSTA	1475	cystatin A	-1.46	3.57E-13	1
VCAN	1462	versican	-1.51	4.25E-11	1
OPRK1	4986	opioid receptor kappa 1	1.30	5.54E-05	1
BIN2	51411	bridging integrator 2	-1.80	2.03E-17	1
SLC27A2	11001	solute carrier family 27 member 2	-1.35	4.57E-08	1
BCL2A1	597	BCL2 related protein A1	-1.67	8.05E-12	1
IGSF6	10261	immunoglobulin superfamily member 6	-1.56	2.43E-14	1
FCN3	8547	ficolin 3	1.15	6.96E-10	1
ACKR3	57007	atypical chemokine receptor 3	1.27	4.44E-10	1
CSRNP1	64651	cysteine and serine rich nuclear protein 1	1.06	4.21E-15	1
TIMD4	91937	T cell immunoglobulin and mucin domain containing 4	-1.13	1.06E-11	1
PLA2G7	7941	phospholipase A2 group VII	-2.82	1.57E-13	1
IGFBP1	3484	insulin like growth factor binding protein 1	-3.35	3.68E-05	1
CRH	1392	corticotropin releasing hormone	6.14	1.72E-24	1
GDPD1	284161	glycerophosphodiester phosphodiesterase domain containing 1	1.56	2.08E-09	1
SST	6750	somatostatin	-1.01	1.42E-07	1
STAT5B	6777	signal transducer and activator of transcription 5B	1.53	5.78E-13	1
VCAM1	7412	vascular cell adhesion molecule 1	-1.20	4.30E-08	1
HPGDS	27306	hematopoietic prostaglandin D synthase	-1.29	1.81E-10	1
ALB	213	albumin	-1.94	3.05E-08	1
CXCL3	2921	C-X-C motif chemokine ligand 3	-1.01	6.38E-07	1
CXCL5	6374	C-X-C motif chemokine ligand 5	-1.56	9.99E-07	1
PPBP	5473	pro-platelet basic protein	-1.42	9.88E-08	1
PF4	5196	platelet factor 4	-1.64	4.82E-16	1
CLEC3B	7123	C-type lectin domain family 3 member B	1.01	4.05E-07	1
CCR1	1230	C-C motif chemokine receptor 1	-1.03	7.44E-08	1
LAIR2	3904	leukocyte associated immunoglobulin like receptor 2	-2.25	5.66E-07	1
ANGPTL4	51129	angiopoietin like 4	1.24	7.66E-10	1
PRL	5617	prolactin	-1.35	1.81E-02	1
ADORA2B	136	adenosine A2b receptor	-1.26	1.73E-11	1
CXCL8	3576	C-X-C motif chemokine ligand 8	-2.46	4.21E-06	1
OLR1	4973	oxidized low density lipoprotein receptor 1	-1.01	7.66E-04	1
CSHL1	1444	chorionic somatomammotropin hormone like 1	2.53	1.89E-14	1
SLC5A6	8884	solute carrier family 5 member 6	1.02	2.02E-05	1
PRG2	5553	proteoglycan 2, pro eosinophil major basic protein	-1.80	1.48E-05	1
LEP	3952	leptin	2.58	6.52E-06	1
GH1	2688	growth hormone 1	2.33	9.87E-13	1
C7	730	complement C7	1.07	1.54E-03	1
NFKBIZ	64332	NFKB inhibitor zeta	-1.12	7.05E-06	1
IL1R2	7850	interleukin 1 receptor type 2	-1.04	1.94E-02	1
CFD	1675	complement factor D	-1.05	2.13E-08	1
GH2	2689	growth hormone 2	4.71	2.44E-15	1
FGG	2266	fibrinogen gamma chain	-1.07	7.09E-04	1

CXCL14	9547	C-X-C motif chemokine ligand 14	-2.19	1.25E-16	1
C15ORF48	84419	chromosome 15 open reading frame 48	-1.29	9.95E-06	1
DHRS2	10202	dehydrogenase/reductase 2	1.07	2.20E-04	1
MAFK	7975	MAF bZIP transcription factor K	1.33	5.10E-14	1
MAFF	23764	MAF bZIP transcription factor F	1.66	2.05E-14	1
SLC38A6	145389	solute carrier family 38 member 6	-1.22	2.56E-15	1
SCARA5	286133	scavenger receptor class A member 5	-1.06	1.64E-02	1
SASH3	54440	SAM and SH3 domain containing 3	-1.09	9.29E-20	1
CCL20	6364	C-C motif chemokine ligand 20	-2.10	3.39E-05	1
G0S2	50486	G0/G1 switch 2	-1.08	2.32E-09	1
CFH	3075	complement factor H	1.64	3.18E-10	1
APOA2	336	apolipoprotein A2	-1.63	3.33E-07	1
CD5L	922	CD5 molecule like	-1.58	7.93E-17	1
S100A4	6275	S100 calcium binding protein A4	-1.21	9.43E-05	1
ENTPD1	953	ectonucleoside triphosphate diphosphohydrolase 1	1.74	1.16E-10	1
RBP4	5950	retinol binding protein 4	-1.00	6.98E-03	1
CHRD1	91851	chordin like 1	-1.20	1.30E-02	1
PLAC9	219348	placenta associated 9	-1.42	7.38E-14	1
PLAU	5328	plasminogen activator, urokinase	1.21	6.67E-09	1
CSF3R	1441	colony stimulating factor 3 receptor	1.06	6.12E-05	1
TREM2	54209	triggering receptor expressed on myeloid cells 2	-1.02	7.00E-11	1
SLC30A2	7780	solute carrier family 30 member 2	1.22	6.51E-09	1
BAMBI	25805	BMP and activin membrane bound inhibitor	-1.01	1.21E-08	1
HBE1	3046	hemoglobin subunit epsilon 1	-3.73	1.25E-30	1
HBD	3045	hemoglobin subunit delta	1.29	1.17E-05	1
CD274	29126	CD274 molecule	1.63	1.05E-07	1
GOLM1	51280	golgi membrane protein 1	-1.21	4.64E-18	1
SPP1	6696	secreted phosphoprotein 1	-1.10	1.27E-05	1
AFP	174	alpha fetoprotein	-1.60	7.94E-07	1
OR51E1	143503	olfactory receptor family 51 subfamily E member 1	1.24	2.95E-10	1
RGN	9104	regucalcin	1.25	3.77E-15	1
FGL1	2267	fibrinogen like 1	-1.02	7.22E-12	1
PLA2G2A	5320	phospholipase A2 group IIA	-2.21	1.25E-10	1
IL6	3569	interleukin 6	-1.72	2.02E-05	1
TF	7018	transferrin	-1.19	2.47E-07	1
CXCR4	7852	C-X-C motif chemokine receptor 4	1.02	7.80E-08	1
NOTUM	147111	notum, palmitoleoyl-protein carboxylesterase	-1.59	6.90E-04	1
AHSG	197	alpha 2-HS glycoprotein	-1.05	2.33E-06	1
SNX10	29887	sorting nexin 10	1.07	2.50E-08	1
SELENBP1	8991	selenium binding protein 1	-1.51	1.47E-12	1
SERPINB2	5055	serpin family B member 2	1.79	8.97E-07	1
COLEC11	78989	collectin subfamily member 11	-1.07	5.76E-11	1
PAEP	5047	progesterone associated endometrial protein	-3.94	1.37E-09	1
TREML2	79865	triggering receptor expressed on myeloid cells like 2	2.20	2.17E-10	1
RYBP	23429	RING1 and YY1 binding protein	1.20	3.05E-07	1
TMEM158	25907	transmembrane protein 158 (gene/pseudogene)	-1.18	2.13E-11	1

GRAMD2B	65983	GRAM domain containing 2B	1.48	1.12E-15	1
CXCL2	2920	C-X-C motif chemokine ligand 2	-1.84	2.15E-06	1
DIO3	1735	iodothyronine deiodinase 3	-1.30	3.17E-18	1
ADM	133	adrenomedullin	1.32	7.58E-09	1
SLC35F2	54733	solute carrier family 35 member F2	1.22	2.02E-12	1
IDO2	169355	indoleamine 2,3-dioxygenase 2	2.50	3.63E-14	1
TDO2	6999	tryptophan 2,3-dioxygenase	-1.83	5.19E-13	1
MMP12	4321	matrix metalloproteinase 12	-2.62	3.82E-07	1
DPEP2	64174	dipeptidase 2	1.39	1.68E-12	1
APOC1	341	apolipoprotein C1	-1.56	1.92E-12	1
FAM20A	54757	FAM20A golgi associated secretory pathway pseudokinase	-1.05	7.67E-15	1
VEGFA	7422	vascular endothelial growth factor A	-1.08	1.54E-04	1
TCHH	7062	trichohyalin	3.63	8.57E-23	1
CGA	1081	glycoprotein hormones, alpha polypeptide	1.05	3.58E-04	1
CCL4	6351	C-C motif chemokine ligand 4	-2.01	3.36E-10	1
RIPOR2	9750	RHO family interacting cell polarization regulator 2	1.50	4.98E-11	1
CYP51A1	1595	cytochrome P450 family 51 subfamily A member 1	-1.14	1.07E-15	2
SLC7A2	6542	solute carrier family 7 member 2	1.97	1.05E-09	2
PDK4	5166	pyruvate dehydrogenase kinase 4	1.10	5.79E-06	2
SULT2B1	6820	sulfotransferase family 2B member 1	1.04	3.03E-09	2
CYP24A1	1591	cytochrome P450 family 24 subfamily A member 1	-1.46	4.24E-17	2
STS	412	steroid sulfatase	1.35	5.31E-04	2
TMEM97	27346	transmembrane protein 97	-1.50	6.65E-19	2
ODC1	4953	ornithine decarboxylase 1	-1.02	3.57E-06	2
DUSP4	1846	dual specificity phosphatase 4	2.05	1.75E-13	2
SOX4	6659	SRY-box transcription factor 4	-1.08	1.89E-14	2
DLGAP5	9787	DLG associated protein 5	-1.29	2.55E-13	2
GINS2	51659	GINS complex subunit 2	-1.14	7.03E-13	2
GGH	8836	gamma-glutamyl hydrolase	-1.31	1.10E-08	2
FDX1	2230	ferredoxin 1	1.23	2.55E-08	2
KIF11	3832	kinesin family member 11	-1.19	1.18E-15	2
MSMO1	6307	methylsterol monooxygenase 1	-1.67	3.83E-13	2
RORA	6095	RAR related orphan receptor A	1.19	2.89E-18	2
SLC7A5	8140	solute carrier family 7 member 5	1.30	7.90E-08	2
ALDH2	217	aldehyde dehydrogenase 2 family member	-1.06	1.51E-13	2
TIPIN	54962	TIMELESS interacting protein	-1.02	2.77E-13	2
GINS1	9837	GINS complex subunit 1	-1.40	9.29E-20	2
MXD1	4084	MAX dimerization protein 1	1.51	9.28E-15	2
DEPDC1B	55789	DEP domain containing 1B	2.21	2.62E-06	2
PAPSS1	9061	3'-phosphoadenosine 5'-phosphosulfate synthase 1	-1.02	6.13E-09	2
DBF4	10926	DBF4 zinc finger	-1.11	1.56E-16	2
ANLN	54443	anillin actin binding protein	-1.13	5.84E-08	2
SQLE	6713	squalene epoxidase	-1.09	9.68E-11	2

CYP11A1	1583	cytochrome P450 family 11 subfamily A member 1	2.02	2.39E-09	2
SAT2	112483	spermidine/spermine N1-acetyltransferase family member 2	-1.01	1.26E-10	2
MISP3	113230	MISP family member 3	1.16	9.18E-15	2
PLA1A	51365	phospholipase A1 member A	1.26	1.38E-20	2
FADS2	9415	fatty acid desaturase 2	-1.17	1.55E-14	2
ZNF117	51351	zinc finger protein 117	1.62	3.00E-08	2
HMGB2	3148	high mobility group box 2	-1.03	3.08E-08	2
RPS23	6228	ribosomal protein S23	-1.15	2.04E-10	2
NOS3	4846	nitric oxide synthase 3	1.76	5.21E-10	2
ALDH1A1	216	aldehyde dehydrogenase 1 family member A1	-1.91	4.14E-08	2
PCLAF	9768	PCNA clamp associated factor	-1.49	2.47E-17	2
ZNF91	7644	zinc finger protein 91	1.02	1.58E-06	2
SLC39A4	55630	solute carrier family 39 member 4	-1.22	1.30E-21	2
HLTF	6596	helicase like transcription factor	-1.01	1.21E-08	2
RPL4	6124	ribosomal protein L4	-1.35	6.26E-17	2
YPEL2	388403	yippee like 2	1.10	1.20E-18	2
IFT80	57560	intraflagellar transport 80	-1.06	6.92E-14	2
METTL1	4234	methyltransferase like 1	-1.23	8.09E-14	2
CCNG2	901	cyclin G2	-1.07	3.57E-07	2
HSD11B2	3291	hydroxysteroid 11-beta dehydrogenase 2	2.97	6.04E-14	2
EZH2	2146	enhancer of zeste 2 polycomb repressive complex 2 subunit	-1.02	1.32E-17	2
ZNF554	115196	zinc finger protein 554	1.50	6.12E-10	2
FADS1	3992	fatty acid desaturase 1	-1.36	2.98E-15	2
HMGCS1	3157	3-hydroxy-3-methylglutaryl-CoA synthase 1	-1.03	3.80E-11	2
DNMT3B	1789	DNA methyltransferase 3 beta	-1.04	5.06E-19	2
CDKN3	1033	cyclin dependent kinase inhibitor 3	-1.24	1.87E-13	2
ZBTB16	7704	zinc finger and BTB domain containing 16	1.32	2.38E-09	2
MAOA	4128	monoamine oxidase A	1.54	2.46E-08	2
ZNF600	162966	zinc finger protein 600	1.16	3.77E-11	2
RPL22L1	200916	ribosomal protein L22 like 1	-1.13	1.82E-16	2
MORN3	283385	MORN repeat containing 3	1.46	7.23E-11	2
UBE2C	11065	ubiquitin conjugating enzyme E2 C	-1.13	3.95E-12	2
ICK	22858	intestinal cell kinase	1.26	3.96E-09	2
MAGEA4	4103	MAGE family member A4	-1.14	3.63E-07	2
PLAGL1	5325	PLAG1 like zinc finger 1	-1.03	1.07E-08	2
RRM2	6241	ribonucleotide reductase regulatory subunit M2	-1.74	3.84E-18	2
PPP1R14C	81706	protein phosphatase 1 regulatory inhibitor subunit 14C	2.33	6.40E-21	2
KMO	8564	kynurenine 3-monooxygenase	1.33	2.58E-05	2
ACAT2	39	acetyl-CoA acetyltransferase 2	-1.57	8.54E-18	2
NME7	29922	NME/NM23 family member 7	-1.01	1.16E-07	2
PHGDH	26227	phosphoglycerate dehydrogenase	-1.08	8.74E-12	2
HSD3B1	3283	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	2.28	1.64E-07	2
TTK	7272	TTK protein kinase	-1.06	1.64E-11	2

BAG2	9532	BCL2 associated athanogene 2	-1.23	1.04E-15	2
ITGB3BP	23421	integrin subunit beta 3 binding protein	-1.03	1.09E-08	2
LCOR	84458	ligand dependent nuclear receptor corepressor	1.35	1.14E-13	2
DHCR24	1718	24-dehydrocholesterol reductase	-1.24	4.22E-07	2
CEP55	55165	centrosomal protein 55	-1.17	1.44E-15	2
ASS1	445	argininosuccinate synthase 1	-1.47	4.87E-13	2
CDC20	991	cell division cycle 20	-1.13	4.59E-11	2
NDUFS5	4725	NADH:ubiquinone oxidoreductase subunit S5	-1.33	8.54E-15	2
PPIL1	51645	peptidylprolyl isomerase like 1	-1.07	3.44E-20	2
ZWINT	11130	ZW10 interacting kinetochore protein	-1.21	1.50E-15	2
PAQR7	164091	progesterin and adipoQ receptor family member 7	1.04	7.56E-12	2
KAZN	23254	kazrin, periplakin interacting protein	1.01	3.48E-12	2
ECI2	10455	enoyl-CoA delta isomerase 2	-1.03	9.65E-11	2
SLC22A18	5002	solute carrier family 22 member 18	-1.25	3.29E-14	2
SLC7A4	6545	solute carrier family 7 member 4	1.05	3.39E-06	2
SNRNP25	79622	small nuclear ribonucleoprotein U11/U12 subunit 25	-1.10	2.20E-14	2
BAZ2B	29994	bromodomain adjacent to zinc finger domain 2B	1.01	5.13E-10	2
SYT12	91683	synaptotagmin 12	1.14	2.43E-15	2
MTHFD2	10797	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	-1.47	1.49E-18	2
METTL7B	196410	methyltransferase like 7B	-1.10	4.63E-10	2
KIF20A	10112	kinesin family member 20A	-1.21	4.17E-13	2
SULT1A1	6817	sulfotransferase family 1A member 1	-1.61	2.27E-11	2
MGST1	4257	microsomal glutathione S-transferase 1	-1.32	5.67E-04	2
CYP19A1	1588	cytochrome P450 family 19 subfamily A member 1	2.42	1.45E-08	2
ADHFE1	137872	alcohol dehydrogenase iron containing 1	2.03	2.25E-17	2
GATM	2628	glycine amidinotransferase	-1.04	2.62E-08	2
CENPK	64105	centromere protein K	-1.23	2.20E-14	2
PRPS2	5634	phosphoribosyl pyrophosphate synthetase 2	-1.04	9.23E-09	2
SPTLC3	55304	serine palmitoyltransferase long chain base subunit 3	2.24	2.55E-12	2
KIF2A	3796	kinesin family member 2A	1.07	7.65E-08	2
FAH	2184	fumarylacetoacetate hydrolase	-1.04	1.54E-11	2
TC2N	123036	tandem C2 domains, nuclear	1.45	2.10E-08	2
RDH13	112724	retinol dehydrogenase 13	1.35	2.14E-12	2
PAQR8	85315	progesterin and adipoQ receptor family member 8	-1.22	1.32E-16	2
TOP2A	7153	DNA topoisomerase II alpha	-1.38	6.08E-11	2
AOC1	26	amine oxidase copper containing 1	-1.16	5.90E-03	2
SMYD3	64754	SET and MYND domain containing 3	-1.05	2.77E-12	2
KLHL5	51088	kelch like family member 5	1.03	2.03E-05	2
PBK	55872	PDZ binding kinase	-1.22	1.30E-12	2
HS6ST2	90161	heparan sulfate 6-O-sulfotransferase 2	-1.11	2.68E-09	2
KPNA2	3838	karyopherin subunit alpha 2	-1.21	3.54E-15	2
HSD3B2	3284	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2	-1.71	1.64E-17	2

SLC38A1	81539	solute carrier family 38 member 1	1.15	5.01E-07	2
NUSAP1	51203	nucleolar and spindle associated protein 1	-1.26	1.89E-14	2
HSD17B1	3292	hydroxysteroid 17-beta dehydrogenase 1	1.22	5.72E-05	2
ZNF468	90333	zinc finger protein 468	1.37	1.13E-07	2
TMSB15A	11013	thymosin beta 15a	-1.41	1.44E-14	2
FMO1	2326	flavin containing dimethylaniline monooxygenase 1	-1.51	2.25E-17	2
TPTE	7179	transmembrane phosphatase with tensin homology	-1.02	5.39E-06	2
KRT14	3861	keratin 14	-1.39	2.68E-03	3
EPS8L1	54869	EPS8 like 1	1.94	3.63E-15	3
PHYH	5264	phytanoyl-CoA 2-hydroxylase	-1.07	2.34E-15	3
ESAM	90952	endothelial cell adhesion molecule	1.31	4.28E-05	3
TRIM62	55223	tripartite motif containing 62	1.06	2.92E-10	3
S100P	6286	S100 calcium binding protein P	1.16	2.50E-07	3
CREBRF	153222	CREB3 regulatory factor	1.33	7.45E-14	3
CRYBG2	55057	crystallin beta-gamma domain containing 2	1.04	1.38E-12	3
UCP2	7351	uncoupling protein 2	-1.33	2.45E-09	3
AGPAT4	56895	1-acylglycerol-3-phosphate O-acyltransferase 4	-1.03	1.18E-16	3
MTMR4	9110	myotubularin related protein 4	1.51	9.96E-11	3
SFN	2810	stratifin	-1.38	2.74E-05	3
NCCRP1	342897	NCCRP1, F-box associated domain containing	2.48	4.30E-17	3
GALNT15	117248	polypeptide N-acetylgalactosaminyltransferase 15	1.82	1.37E-11	3
SCNN1B	6338	sodium channel epithelial 1 beta subunit	2.32	2.56E-15	3
GUCA2A	2980	guanylate cyclase activator 2A	1.12	4.07E-06	3
EMP2	2013	epithelial membrane protein 2	1.36	2.55E-12	3
NECTIN4	81607	nectin cell adhesion molecule 4	1.63	2.02E-15	3
OLAH	55301	oleoyl-ACP hydrolase	2.54	6.15E-08	3
LAD1	3898	ladinin 1	1.25	1.92E-12	3
ALPP	250	alkaline phosphatase, placental	6.65	7.62E-30	3
GALNT3	2591	polypeptide N-acetylgalactosaminyltransferase 3	-1.13	8.89E-07	3
TRIM40	135644	tripartite motif containing 40	1.25	1.58E-06	3
MUC16	94025	mucin 16, cell surface associated	-1.01	1.33E-02	3
CLDN8	9073	claudin 8	1.28	4.04E-07	3
AGR2	10551	anterior gradient 2, protein disulphide isomerase family member	-1.11	1.42E-09	3
ATP2B1	490	ATPase plasma membrane Ca <sup>2+</sup> transporting 1	-1.27	7.01E-19	3
IP6K3	117283	inositol hexakisphosphate kinase 3	1.03	1.15E-15	3
HLA-G	3135	major histocompatibility complex, class I, G	-2.13	3.75E-09	3
MUC20	200958	mucin 20, cell surface associated	1.24	1.82E-04	3
NECTIN3	25945	nectin cell adhesion molecule 3	2.60	3.74E-16	3
ELOVL7	79993	ELOVL fatty acid elongase 7	1.86	2.66E-18	3
TOB1	10140	transducer of ERBB2, 1	-1.09	5.80E-05	3
ST3GAL4	6484	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	1.28	7.21E-11	3
CYB5R2	51700	cytochrome b5 reductase 2	-1.18	3.16E-13	3
IFI27	3429	interferon alpha inducible protein 27	1.14	3.17E-06	3
ITGA3	3675	integrin subunit alpha 3	1.29	6.28E-10	4
MMP11	4320	matrix metalloproteinase 11	2.44	1.77E-14	4
NID2	22795	nidogen 2	-1.16	1.10E-10	4
MMP2	4313	matrix metalloproteinase 2	-1.02	6.71E-06	4

SFRP1	6422	secreted frizzled related protein 1	-1.21	1.05E-05	4
PLAT	5327	plasminogen activator, tissue type	1.01	9.67E-05	4
PCOLCE	5118	procollagen C-endopeptidase enhancer	-1.82	1.09E-17	4
CAP2	10486	cyclase associated actin cytoskeleton regulatory protein 2	1.43	2.74E-07	4
ISL1	3670	ISL LIM homeobox 1	-1.17	1.05E-07	4
LOX	4015	lysyl oxidase	-1.03	6.17E-03	4
HAND1	9421	heart and neural crest derivatives expressed 1	-1.09	6.01E-05	4
HES1	3280	hes family bHLH transcription factor 1	1.20	7.38E-08	4
INHBA	3624	inhibin subunit beta A	3.58	1.30E-15	4
TWIST1	7291	twist family bHLH transcription factor 1	1.06	3.74E-08	4
INHA	3623	inhibin subunit alpha	1.44	2.66E-10	4
COL21A1	81578	collagen type XXI alpha 1 chain	-1.50	4.85E-13	4
CNN1	1264	calponin 1	1.00	1.80E-04	4
GDF15	9518	growth differentiation factor 15	2.98	6.80E-14	4
LGALS3	3958	galectin 3	-1.30	9.74E-07	4
RAB11FIP5	26056	RAB11 family interacting protein 5	1.17	4.18E-09	4
GCM1	8521	glial cells missing transcription factor 1	1.24	3.57E-04	4
EPYC	1833	epiphycan	-1.33	4.01E-03	4
CASQ2	845	calsequestrin 2	1.05	3.31E-09	4
FLT4	2324	fms related tyrosine kinase 4	1.20	3.00E-06	4
DOK5	55816	docking protein 5	-1.07	2.47E-17	4
TIMP2	7077	TIMP metalloproteinase inhibitor 2	1.04	7.28E-07	4
PDZRN3	23024	PDZ domain containing ring finger 3	-1.15	1.19E-16	4
TXK	7294	TXK tyrosine kinase	2.69	2.79E-21	4
PCDH10	57575	protocadherin 10	-1.45	2.86E-14	4
WNT5A	7474	Wnt family member 5A	-1.20	2.78E-05	4
WNT2	7472	Wnt family member 2	-1.92	1.09E-12	4
AMIGO2	347902	adhesion molecule with Ig like domain 2	-1.48	8.06E-21	4
CDH11	1009	cadherin 11	-1.43	9.25E-14	4
CDH13	1012	cadherin 13	-1.18	2.52E-12	4
GATA6	2627	GATA binding protein 6	-1.55	4.44E-22	4
REN	5972	renin	-2.68	6.04E-25	4
OSR1	130497	odd-skipped related transcription factor 1	-1.52	1.45E-10	4
HAPLN1	1404	hyaluronan and proteoglycan link protein 1	-2.36	9.25E-17	4
ARHGAP26	23092	Rho GTPase activating protein 26	1.24	3.93E-08	4
TNNT3	7140	troponin T3, fast skeletal type	2.13	1.54E-13	4
ANK3	288	ankyrin 3	1.67	6.18E-10	4
PLOD2	5352	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	-1.40	3.01E-05	4
BMP6	654	bone morphogenetic protein 6	1.53	8.20E-17	4
ADGRF5	221395	adhesion G protein-coupled receptor F5	1.16	3.47E-07	4
THY1	7070	Thy-1 cell surface antigen	-1.04	1.64E-05	4
ABI3BP	25890	ABI family member 3 binding protein	-2.07	2.03E-06	4
ADAMTS5	11096	ADAM metalloproteinase with thrombospondin type 1 motif 5	-1.51	2.24E-11	4
DKK2	27123	dickkopf WNT signaling pathway inhibitor 2	-1.77	8.63E-19	4
ADAMTS3	9508	ADAM metalloproteinase with thrombospondin type 1 motif 3	1.85	1.64E-13	4

PCDH1	5097	protocadherin 1	1.04	3.48E-09	4
TIMP4	7079	TIMP metalloproteinase inhibitor 4	1.09	6.35E-14	4
PRKAG2	51422	protein kinase AMP-activated non-catalytic subunit gamma 2	1.02	1.72E-08	4
ARHGAP17	55114	Rho GTPase activating protein 17	1.43	2.90E-11	4
STC1	6781	stanniocalcin 1	-2.14	3.61E-10	4
PKDCC	91461	protein kinase domain containing, cytoplasmic	-1.38	1.20E-18	4
NIPAL1	152519	NIPA like domain containing 1	1.81	1.96E-11	4
COL6A3	1293	collagen type VI alpha 3 chain	-1.48	5.27E-12	4
TLR3	7098	toll like receptor 3	1.09	1.16E-03	4
STEAP1	26872	STEAP family member 1	-1.13	1.81E-07	4
FNDC1	84624	fibronectin type III domain containing 1	-1.79	2.46E-23	4
COL14A1	7373	collagen type XIV alpha 1 chain	-1.07	2.21E-05	4
COL6A2	1292	collagen type VI alpha 2 chain	-1.01	9.63E-11	4
IGFBP6	3489	insulin like growth factor binding protein 6	-1.41	5.69E-05	4
INSR	3643	insulin receptor	1.53	6.62E-13	4
PITX2	5308	paired like homeodomain 2	-1.52	1.18E-17	4
BMP1	649	bone morphogenetic protein 1	1.36	1.42E-10	4
SYNPO2	171024	synaptopodin 2	1.23	3.39E-08	4
PAM	5066	peptidylglycine alpha-amidating monooxygenase	1.49	4.11E-12	4
HAS2	3037	hyaluronan synthase 2	-1.48	3.80E-09	4
MRC2	9902	mannose receptor C type 2	-1.25	1.70E-19	4
CD248	57124	CD248 molecule	-1.11	1.68E-04	4
RASGRP1	10125	RAS guanyl releasing protein 1	2.41	9.40E-18	4
SLCO2A1	6578	solute carrier organic anion transporter family member 2A1	1.47	5.83E-06	4
LIMCH1	22998	LIM and calponin homology domains 1	2.17	2.09E-14	4
CAPN6	827	calpain 6	2.44	5.83E-11	4
RAP2B	5912	RAP2B, member of RAS oncogene family	-1.27	2.17E-25	4
FIBIN	387758	fin bud initiation factor homolog	-1.30	1.70E-15	4
OLFML3	56944	olfactomedin like 3	-1.33	4.40E-14	4
LRCH2	57631	leucine rich repeats and calponin homology domain containing 2	-1.26	2.12E-12	4
ARHGEF37	389337	Rho guanine nucleotide exchange factor 37	1.02	8.60E-11	4
NOG	9241	noggin	-1.58	3.10E-07	4
PAPPA	5069	pappalysin 1	2.70	2.20E-12	4
FBLN1	2192	fibulin 1	1.23	7.06E-05	4
IRX3	79191	iroquois homeobox 3	-1.08	3.19E-12	4
CRIP1	1396	cysteine rich protein 1	1.24	6.55E-08	4
OLFML1	283298	olfactomedin like 1	-1.01	4.10E-14	4
IGDCC3	9543	immunoglobulin superfamily DCC subclass member 3	-1.16	1.20E-14	4
FLRT2	23768	fibronectin leucine rich transmembrane protein 2	-1.16	2.52E-14	4
ABCC5	10057	ATP binding cassette subfamily C member 5	-1.03	9.40E-18	4
HOMER1	9456	homer scaffold protein 1	-1.16	1.31E-17	4
PTPRB	5787	protein tyrosine phosphatase receptor type B	1.16	1.04E-07	4
GLDN	342035	gliomedin	2.67	5.39E-12	4
SYNM	23336	synemin	1.03	3.28E-10	4

HEY1	23462	hes related family bHLH transcription factor with YRPW motif 1	1.04	3.60E-08	4
STX3	6809	syntaxin 3	1.32	2.33E-14	4
SYDE1	85360	synapse defective Rho GTPase homolog 1	1.24	9.17E-11	4
TNNI3	7137	troponin I3, cardiac type	-1.31	5.70E-24	4
PRELP	5549	proline and arginine rich end leucine rich repeat protein	1.06	1.22E-14	4
POMGNT2	84892	protein O-linked mannose N-acetylglucosaminyltransferase 2 (beta 1,4-)	-1.20	7.00E-19	4
LRRC17	10234	leucine rich repeat containing 17	-1.06	1.30E-03	4
CORO6	84940	coronin 6	2.13	5.03E-09	4
SPRY4	81848	sprouty RTK signaling antagonist 4	1.37	2.04E-10	4
RSPO3	84870	R-spondin 3	-1.16	4.86E-07	4
ANK2	287	ankyrin 2	-1.04	6.05E-14	4
MAP1LC3C	440738	microtubule associated protein 1 light chain 3 gamma	1.22	1.66E-11	4
COL6A6	131873	collagen type VI alpha 6 chain	-1.38	2.34E-23	4
ADGRG6	57211	adhesion G protein-coupled receptor G6	1.93	3.15E-10	4
ADAM12	8038	ADAM metallopeptidase domain 12	1.65	1.25E-07	4
HTRA1	5654	HtrA serine peptidase 1	1.82	2.48E-10	4
LRRC1	55227	leucine rich repeat containing 1	1.23	2.85E-12	4
KANK4	163782	KN motif and ankyrin repeat domains 4	-1.07	1.73E-14	4
COL5A1	1289	collagen type V alpha 1 chain	-1.22	9.09E-10	4
TNMD	64102	tenomodulin	1.25	4.10E-18	4
RAB42	115273	RAB42, member RAS oncogene family	-1.09	5.25E-13	4
DKK1	22943	dickkopf WNT signaling pathway inhibitor 1	-2.50	1.80E-06	4
RND3	390	Rho family GTPase 3	-1.30	2.78E-06	4
ITGBL1	9358	integrin subunit beta like 1	-1.02	2.29E-09	4
HES2	54626	hes family bHLH transcription factor 2	1.29	8.03E-17	4
SPIRE2	84501	spire type actin nucleation factor 2	1.31	3.19E-10	4
DNAAF1	123872	dynein axonemal assembly factor 1	1.22	2.89E-10	4
BMP2	650	bone morphogenetic protein 2	-1.22	3.89E-12	4
TFAP2A	7020	transcription factor AP-2 alpha	1.07	1.65E-08	4
RHOBTB3	22836	Rho related BTB domain containing 3	1.47	4.73E-08	4
IGSF5	150084	immunoglobulin superfamily member 5	1.16	5.58E-05	4
CPXM1	56265	carboxypeptidase X, M14 family member 1	-1.50	7.19E-08	4
SDC1	6382	syndecan 1	3.21	1.74E-18	4
SMTNL2	342527	smoothelin like 2	-1.39	2.07E-23	4
PDGFD	80310	platelet derived growth factor D	1.46	3.39E-09	4
PTGFRN	5738	prostaglandin F2 receptor inhibitor	1.37	6.53E-13	4
FHL1	2273	four and a half LIM domains 1	1.42	3.48E-07	4
MFAP4	4239	microfibril associated protein 4	-1.38	1.09E-07	4
MYH11	4629	myosin heavy chain 11	1.81	3.20E-11	4
SNAI2	6591	snail family transcriptional repressor 2	-1.68	1.14E-09	4
VGLL3	389136	vestigial like family member 3	1.71	4.58E-15	4
PDZRN4	29951	PDZ domain containing ring finger 4	-1.01	1.32E-10	4
MDK	4192	midkine	-1.79	8.06E-21	4
DAPK1	1612	death associated protein kinase 1	1.29	7.18E-12	4

GULP1	51454	GULP PTB domain containing engulfment adaptor 1	1.61	1.34E-06	4
CHN1	1123	chimerin 1	-1.27	2.50E-09	4
NANOS1	340719	nanos C2HC-type zinc finger 1	-1.08	2.62E-12	4
FGD4	121512	FYVE, RhoGEF and PH domain containing 4	1.11	9.29E-20	4
LRRN3	54674	leucine rich repeat neuronal 3	-1.86	1.53E-15	4
LAMA2	3908	laminin subunit alpha 2	-1.04	1.41E-06	4
PHLDB2	90102	pleckstrin homology like domain family B member 2	1.11	3.57E-06	4
CLMP	79827	CXADR like membrane protein	-1.19	2.25E-14	4
OSR2	116039	odd-skipped related transcription factor 2	-1.17	2.28E-09	4
HS3ST6	64711	heparan sulfate-glucosamine 3-sulfotransferase 6	-1.26	8.15E-16	4
CD200	4345	CD200 molecule	2.40	6.24E-09	4
ROBO1	6091	roundabout guidance receptor 1	-1.32	8.06E-21	4
SLIT2	9353	slit guidance ligand 2	-1.29	2.62E-10	4
GPX8	493869	glutathione peroxidase 8 (putative)	-1.08	1.53E-07	4
CRIP2	1397	cysteine rich protein 2	1.28	4.93E-13	4
RAI14	26064	retinoic acid induced 14	1.11	7.97E-08	4
FHDC1	85462	FH2 domain containing 1	1.65	2.99E-09	4
RAP1GAP	5909	RAP1 GTPase activating protein	1.45	6.76E-14	4
ZNRF3	84133	zinc and ring finger 3	-1.11	2.64E-13	4
HOPX	84525	HOP homeobox	2.27	6.31E-13	4
PGF	5228	placental growth factor	1.56	9.05E-06	4
SPON1	10418	spondin 1	-1.05	1.20E-07	4
TNNT1	7138	troponin T1, slow skeletal type	-1.65	3.31E-11	4
ABCB1	5243	ATP binding cassette subfamily B member 1	1.21	1.01E-03	4
PLCL2	23228	phospholipase C like 2	1.80	8.29E-16	4
LRRC8E	80131	leucine rich repeat containing 8 VRAC subunit E	1.36	5.56E-13	4
LRRC39	127495	leucine rich repeat containing 39	-1.21	2.56E-15	4
CRABP2	1382	cellular retinoic acid binding protein 2	-1.53	2.43E-10	4
SLC2A11	66035	solute carrier family 2 member 11	2.34	1.03E-14	4
FURIN	5045	furin, paired basic amino acid cleaving enzyme	1.38	9.60E-09	4
SEMA3B	7869	semaphorin 3B	2.98	6.12E-16	4
CKLF	51192	chemokine like factor	-1.03	5.86E-07	5
SLC28A2	9153	solute carrier family 28 member 2	-1.83	1.20E-18	5
PRTFDC1	56952	phosphoribosyl transferase domain containing 1	-1.35	9.87E-20	5
UPP1	7378	uridine phosphorylase 1	1.36	1.45E-15	5
CMTM7	112616	CKLF like MARVEL transmembrane domain containing 7	-1.04	1.86E-12	5
SLC23A2	9962	solute carrier family 23 member 2	1.48	2.10E-10	5
VWA5A	4013	von Willebrand factor A domain containing 5A	1.19	2.60E-10	5
CLIC5	53405	chloride intracellular channel 5	1.44	2.23E-08	6
ECRG4	84417	ECRG4 augurin precursor	-2.01	5.08E-18	6
TTYH2	94015	tweety family member 2	-1.19	9.71E-20	6
ANO2	57101	anoctamin 2	1.08	1.90E-05	6
CLIC2	1193	chloride intracellular channel 2	1.15	3.71E-09	6
GABRE	2564	gamma-aminobutyric acid type A receptor epsilon subunit	1.29	9.95E-04	6

CLIC3	9022	chloride intracellular channel 3	1.05	1.24E-11	6
GABRP	2568	gamma-aminobutyric acid type A receptor pi subunit	-1.01	8.10E-06	6
PSG9	5678	pregnancy specific beta-1-glycoprotein 9	3.53	5.32E-12	7
PSG1	5669	pregnancy specific beta-1-glycoprotein 1	3.75	4.25E-10	7
PSG11	5680	pregnancy specific beta-1-glycoprotein 11	4.97	3.10E-16	7
PSG3	5671	pregnancy specific beta-1-glycoprotein 3	3.86	1.35E-14	7
PSG4	5672	pregnancy specific beta-1-glycoprotein 4	4.06	1.84E-16	7
PSG2	5670	pregnancy specific beta-1-glycoprotein 2	3.16	2.84E-11	7

**ДОДАТОК Д**

**АСОЦІЙОВАНІ З ПРЕЕКЛАМПСІЄЮ ГЕНИ СЕРЕД ДИФЕРЕНЦІЙНО**

**ЕКСПРЕСОВАНИХ ГЕНІВ МІЖ ПЕРШИМ І ДРУГИМ ТА ДРУГИМ І**

**ТРЕТИМ ТРИМЕСТРАМИ ВАГІТНОСТІ**

SYMBOL	ENTREZID	GENENAME	logFC	adj.P.Val	cluster	Comparison
CCL2	6347	C-C motifchemokineligand 2	1.65	3.87E-04	1	1 2
CD36	948	CD36 molecule	1.72	3.99E-03	1	1 2
CXCL1	2919	C-X-C motifchemokineligand 1	1.53	1.18E-02	1	1 2
CXCL2	2920	C-X-C motifchemokineligand 2	1.37	1.18E-02	1	1 2
CXCL8	3576	C-X-C motifchemokineligand 8	1.74	1.40E-02	1	1 2
CYBB	1536	cytochrome b-245 beta chain	1.47	5.43E-04	1	1 2
LCN2	3934	lipocalin 2	-1.99	8.10E-03	1	1 2
MUC1	4582	mucin 1, cell surface associated	-1.06	1.07E-02	1	1 2
REN	5972	renin	-1.24	6.17E-04	1	1 2
S100A8	6279	S100 calciumbinding protein A8	1.87	9.60E-03	1	1 2
S100A9	6280	S100 calciumbinding protein A9	1.69	1.04E-04	1	1 2
WFDC2	10406	WAP four-disulfidecoredomain 2	-1.87	5.85E-03	1	1 2
ADM	133	adrenomedullin	1.26	2.01E-02	2	1 2
AQP1	358	aquaporin 1 (Coltonbloodgroup)	-1.12	8.27E-05	2	1 2
CLU	1191	clusterin	-1.44	8.08E-03	2	1 2
CP	1356	ceruloplasmin	-2.03	2.24E-02	2	1 2
GADD45A	1647	growth arrest and DNA damageinduciblealpha	-1.38	6.04E-03	2	1_2
GADD45B	4616	growth arrest and DNA damageinducible beta	1.01	8.27E-05	2	1_2
IL6	3569	interleukin 6	1.69	3.36E-03	2	1 2
PLAT	5327	plasminogenactivator, tissuetype	-1.34	4.32E-02	2	1 2
SERPINB2	5055	serpinfamily B member 2	1.78	1.91E-02	2	1 2
LPL	4023	lipoproteinlipase	1.27	7.49E-05	3	1 2
LEP	3952	leptin	-2.68	2.53E-03	4	1 2
ACKR3	57007	atypicalchemokinereceptor 3	1.27	4.44E-10	1	2 3
ADM	133	adrenomedullin	1.32	7.58E-09	1	2 3
AHR	196	arylhydrocarbonreceptor	1.35	3.49E-06	1	2 3
ALB	213	albumin	-1.94	3.05E-08	1	2 3
APOE	348	apolipoprotein E	-1.03	4.54E-08	1	2 3
C3	718	complement C3	-1.24	5.19E-05	1	2 3
CCL2	6347	C-C motifchemokineligand 2	-1.27	1.34E-04	1	2 3
CCL20	6364	C-C motifchemokineligand 20	-2.10	3.39E-05	1	2 3
CXCL2	2920	C-X-C motifchemokineligand 2	-1.84	2.15E-06	1	2 3
CXCL8	3576	C-X-C motifchemokineligand 8	-2.46	4.21E-06	1	2 3
FST	10468	folliclestimulating hormone receptor 1	-1.01	2.67E-07	1	2 3
G0S2	50486	G0/G1 switch 2	-1.08	2.32E-09	1	2 3

<b>IGFBP1</b>	3484	insulin like growth factor binding protein 1	-3.35	3.68E-05	1	2_3
<b>IL6</b>	3569	interleukin 6	-1.72	2.02E-05	1	2_3
<b>LEP</b>	3952	leptin	2.58	6.52E-06	1	2_3
<b>LHB</b>	3972	lutinizing hormone subunit beta	1.42	3.86E-05	1	2_3
<b>PRL</b>	5617	prolactin	-1.35	1.81E-02	1	2_3
<b>SELENBP1</b>	8991	selenium binding protein 1	-1.51	1.47E-12	1	2_3
<b>SERPINB2</b>	5055	serpin family B member 2	1.79	8.97E-07	1	2_3
<b>SPP1</b>	6696	secreted phosphoprotein 1	-1.10	1.27E-05	1	2_3
<b>TGFB3</b>	7043	transforming growth factor beta 3	-1.45	2.81E-09	1	2_3
<b>TIMP1</b>	7076	TIMP metalloproteinase inhibitor 1	-1.15	1.31E-04	1	2_3
<b>VCAM1</b>	7412	vascular cell adhesion molecule 1	-1.20	4.30E-08	1	2_3
<b>VEGFA</b>	7422	vascular endothelial growth factor A	-1.08	1.54E-04	1	2_3
ALDH1A1	216	aldehyde dehydrogenase family member A1	-1.91	4.14E-08	2	2_3
CYP11A1	1583	cytochrome P450 family 11 subfamily A member 1	2.02	2.39E-09	2	2_3
NOS3	4846	nitric oxide synthase 3	1.76	5.21E-10	2	2_3
ODC1	4953	ornithine decarboxylase 1	-1.02	3.57E-06	2	2_3
ABCB1	5243	ATP binding cassette subfamily B member 1	1.21	1.01E-03	4	2_3
HOMER1	9456	homer scaffold protein 1	-1.16	1.31E-17	4	2_3
<b>INHBA</b>	3624	Inhibin subunit beta A	3.58	1.30E-15	4	2_3
<b>ITGA3</b>	3675	Integrin subunit alpha 3	1.29	6.28E-10	4	2_3
<b>MMP2</b>	4313	matrix metalloproteinase 2	-1.02	6.71E-06	4	2_3
<b>PAM</b>	5066	Peptidylglycine alpha-amidating monooxygenase	1.49	4.11E-12	4	2_3
<b>PLAT</b>	5327	plasminogen activator, tissue type	1.01	9.67E-05	4	2_3
RND3	390	Rho family GTPase 3	-1.30	2.78E-06	4	2_3
<b>SPON1</b>	10418	spondin 1	-1.05	1.20E-07	4	2_3
THY1	7070	Thy-1 cell surface antigen	-1.04	1.64E-05	4	2_3

Примітка. Гени, білки яких секретуються, виділені жирним шрифтом.

## ДОДАТОК Е

## ВЗЯТІ ДЛЯ СТВОРЕННЯ СПЕЦІАЛІЗОВАНОЇ БАЗИ ДАНИХ

## ІДЕНТИФІКАТОРИ ARRAYEXPRESS

E-GEOD-74341, E-GEOD-73377, E-GEOD-73375, E-GEOD-73374,  
 E-GEOD-60438, E-MTAB-3265, E-MTAB-3348, E-MTAB-3309,  
 E-GEOD-54400, E-GEOD-59274, E-GEOD-57767, E-GEOD-48424,  
 E-GEOD-57050, E-GEOD-54618, E-GEOD-49343, E-GEOD-38747,  
 E-GEOD-47187, E-GEOD-50783, E-GEOD-41681, E-GEOD-44712,  
 E-GEOD-44711, E-GEOD-44667, E-GEOD-43942, E-GEOD-41336,  
 E-GEOD-41331, E-GEOD-40182, E-GEOD-37901, E-GEOD-36083,  
 E-GEOD-35574, E-GEOD-31679, E-GEOD-30186, E-GEOD-15789,  
 E-GEOD-15787, E-GEOD-22526, E-GEOD-24129, E-GEOD-10588,  
 E-GEOD-13155, E-TABM-682, E-GEOD-14722, E-GEOD-12767,  
 E-GEOD-13475, E-GEOD-12216, E-GEOD-9984, E-GEOD-6573,  
 E-GEOD-4100, E-GEOD-4707, E-MEXP-1050

Excluded datasets:

Could not be processed by Bioservices API: E-GEOD-25906.

Irrelevant result: E-MTAB-3732.

Due to microarray design: E-GEOD-15787 (Capitalbio 22K Human oligo array version 1.0 , Capitalbiomammal) microRNA V3.0E-GEOD-22526 (GynObsLU Human 800 PE-associated cDNA) E-MEXP-1050 ( Affymetrix GeneChip Human GenomeFocus Array [HG-Focus ])